

Transceiver Design For SC-FDE Based MIMO Relay Systems

Peiran Wu, Robert Schober, and Vijay Bhargava

Abstract—In this paper, we propose a joint transceiver design for single-carrier frequency-domain equalization (SC-FDE) based multiple-input multiple-output (MIMO) relay systems. To this end, we first derive the optimal minimum mean-squared error linear and decision-feedback frequency-domain equalization filters at the destination along with the corresponding error covariance matrices at the output of the equalizer. Subsequently, we formulate the source and relay precoding matrix design problem as the minimization of a family of Schur-convex and Schur-concave functions of the mean-squared errors at the output of the equalizer under separate power constraints for the source and the relay. By exploiting properties of the error covariance matrix and results from majorization theory, we derive the optimal structures of the source and relay precoding matrices, which allows us to transform the matrix optimization problem into a scalar power optimization problem. Adopting a high signal-to-noise ratio approximation for the objective function, we obtain the global optimal solution for the power allocation variables. Simulation results illustrate the excellent performance of the proposed system and its superiority compared to conventional orthogonal frequency-division multiplexing based MIMO relay systems.

I. INTRODUCTION

Multiple-input multiple-output (MIMO) relay systems utilizing multiple antennas at the relay node have recently received significant research interest due to their potential to enhance network performance [1]. An important research problem for MIMO relay systems is the design of optimal node processing matrices to improve spectral efficiency and/or error performance through efficient utilization of transmit channel state information (CSIT). For example, assuming availability of CSIT at the source and relay nodes and linear processing at the destination, the source and relay processing matrices were optimized for maximization of the relay channel capacity and minimization of the mean-squared error (MSE) in [2], [3] and [4], [5], respectively. In [6], a general framework for linear transceiver optimization in MIMO relay systems was provided for a large family of objective functions, which includes the capacity maximizing and the MSE minimizing designs as special cases. The extension of the results in [6] to multi-hop MIMO relay systems with linear and decision-feedback equalization receivers was investigated in [7] and [8], respectively. More recently, the design of MIMO relay systems with partial or imperfect CSIT at source and relay was considered in [9], [10].

Existing works on transceiver design for MIMO relay systems are based on the assumption of frequency-nonselective (flat) channels [3], [4], [7], [8], [10] or frequency-selective channels in combination with orthogonal frequency-division multiplexing (OFDM) [2], [6], [9]. Since OFDM decomposes a frequency-selective channel into multiple parallel flat subchannels, the transceiver designs developed for frequency-nonselective channels can be extended to OFDM based MIMO relay systems by solving an additional subcarrier power allocation problem across different subcarriers. However, its large peak-to-average power ratio (PAPR) makes OFDM less appealing for application in the uplink of wireless communication systems. Block based single-carrier transmission with frequency-domain equalization (SC-FDE) is a promising alternative to OFDM due to its comparable implementation complexity and lower PAPR [11], [12]. To the best of the authors' knowledge, the optimization of SC-FDE based MIMO relay systems has not been considered in the literature so far. A key difference between SC-FDE based MIMO systems and MIMO-OFDM systems is that the performance of the former depends on the MSEs of each spatial stream whereas the performance of the latter depends on the subcarrier MSEs. This important difference makes the optimization of SC-FDE based MIMO systems more challenging than the optimization of MIMO-OFDM systems.

In this paper, we make the common assumption of perfect CSI at all nodes [2]–[8] and we propose a joint transceiver design for MIMO relay systems employing either frequency-domain linear equalization (FD-LE) or frequency-domain decision feedback equalization (FD-DFE) at the destination. We optimize the source and relay precoding matrices for minimization of a general function of the MSEs of the spatial streams under separate power constraints for source and relay. Specifically, as objective functions we adopt the arithmetic MSE (AMSE), the geometric MSE (GMSE), and the maximum MSE (maxMSE) [6], [14], which are closely related to channel capacity and error rate performance. For the case of FD-LE, we show that the optimal source and relay precoding matrices have a structure very similar to that of the optimal precoding matrices in MIMO-OFDM relay systems. However, the remaining power allocation problem is significantly different from the power allocation problem for MIMO-OFDM relay systems, especially for the GMSE and maxMSE criteria. For FD-DFE, the considered objective functions cannot be explicitly expressed in terms of the optimization variables and depend on the number of feedback filter taps, which makes a direct solution of the optimization problem challenging. However, we can show that for FD-DFE, the three considered

The authors are with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC Canada, V6T, 1Z4, email: {peiranw, rschober, vijayb}@ece.ubc.ca. This paper will be presented in part at the IEEE Global Communications Conference (Globecom), Anaheim, December 2012.

objective functions are equivalent. Furthermore, we develop an upper bound for the objective function which is independent of number of feedback filter taps and is a comparatively simple function of the optimization variables. Interestingly, this upper bound is shown to be identical to the GMSE objective function for the FD-LE receiver. Consequently, a unified solution for the power allocation problem for both FD-LE and FD-DFE can be obtained, which greatly simplifies the design procedure.

The remainder of this paper is organized as follows. In Section II, the system model is presented. In Section III, the optimal minimum MSE (MMSE) FDE filters and the corresponding stream error covariance (CV) matrices are derived. The optimal source and relay precoding matrices are presented in Section IV. Simulation results are given in Section V, and some conclusions are drawn in Section VI.

In this paper, $\text{tr}(\mathbf{A})$, \mathbf{A}^{-1} , \mathbf{A}^T , and \mathbf{A}^\dagger denote the trace, inverse, transpose, and conjugate transpose of matrix \mathbf{A} , respectively. $\mathbb{C}^{M \times N}$ denotes the space of all $M \times N$ complex matrices and \mathbf{I}_M is the $M \times M$ identity matrix. $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I}_M)$ indicates that $\mathbf{n} \in \mathbb{C}^{M \times 1}$ is a complex Gaussian distributed vector with zero mean and CV matrix $\sigma_n^2 \mathbf{I}_M$. $E[\cdot]$ and \otimes denote statistical expectation and the Kronecker product, respectively. $\text{blkcirc}([\mathbf{A}_1^T, \mathbf{A}_2^T, \dots, \mathbf{A}_M^T]^T)$ and $\text{blkdiag}([\mathbf{A}_1^T, \mathbf{A}_2^T, \dots, \mathbf{A}_M^T]^T)$ denote a block circular matrix and a block diagonal matrix, respectively, formed by the block-wise vector $[\mathbf{A}_1^T, \mathbf{A}_2^T, \dots, \mathbf{A}_M^T]^T$. \mathbb{F}_N denotes the $N \times N$ Discrete Fourier Transform (DFT) matrix, and x^* denotes the optimal value of x .

II. SYSTEM MODEL

We consider a block transmission system with one source node, S , one relay node, R , and one destination node, D , as shown in Fig. 1. The numbers of antennas at S , R , and D are denoted by N_s , N_r , and N_d , respectively. The number of spatial multiplexing data streams is $M \leq \min\{N_s, N_r, N_d\}$. The transmission is organized in two phases. In the first phase, S processes the information symbols and sends them to R . In the second phase, R processes the received signal and retransmits it to D . We assume there is no direct link between S and D due to the large pathloss and/or shadowing.

The transmit signal of each source antenna is prepended by a cyclic prefix (CP), which comprises the last $N_{g,s} \geq L_g$ symbols of the transmitted source signal, where L_g denotes the largest channel impulse response (CIR) length between any S - R antenna pair.¹ Similarly, the transmit signal of each relay antenna is prepended by a CP, which comprises the last $N_{g,r} \geq L_h$ symbols of the transmitted relay signal, where L_h is the largest CIR length between any R - D antenna pair.

A. Precoding at Source and Relay

Let us denote the n th source data symbol vector as $\mathbf{s}_n = [s_n(1), s_n(2), \dots, s_n(M)]^T$, $n = 0, \dots, N_c - 1$, where N_c is the size of the data block, and $s_n(j)$ denotes the n th symbol of the j th data stream, which is drawn from a constellation with variance σ_s^2 . By stacking all \mathbf{s}_n into one vector, we obtain

$\mathbf{s} = [\mathbf{s}_0^T, \dots, \mathbf{s}_{N_c-1}^T]^T \in \mathbb{C}^{MN_c \times 1}$. The received signal at the destination, \mathbf{y} , can be compactly written as

$$\mathbf{y} = \mathbf{H}_t \mathbf{A}_t \mathbf{G}_t \mathbf{P}_t \mathbf{s} + \mathbf{H}_t \mathbf{A}_t \mathbf{v} + \mathbf{u} \quad (1)$$

with block circular matrices

$$\begin{aligned} \mathbf{P}_t &= \text{blkcirc}([\mathbf{P}_{t,0}^T, \dots, \mathbf{P}_{t,N_c-1}^T]^T), \\ \mathbf{G}_t &= \text{blkcirc}([\mathbf{G}_{t,0}^T, \dots, \mathbf{G}_{t,L_g-1}^T, \mathbf{0}_{N_s \times N_r(N_c-L_g)}]^T), \\ \mathbf{A}_t &= \text{blkcirc}([\mathbf{A}_{t,0}^T, \dots, \mathbf{A}_{t,N_c-1}^T]^T), \\ \mathbf{H}_t &= \text{blkcirc}([\mathbf{H}_{t,0}^T, \dots, \mathbf{H}_{t,L_h-1}^T, \mathbf{0}_{N_r \times N_d(N_c-L_h)}]^T), \end{aligned}$$

where $\mathbf{P}_{t,l} \in \mathbb{C}^{N_s \times M}$, $\mathbf{G}_{t,l} \in \mathbb{C}^{N_r \times N_s}$, $\mathbf{A}_{t,l} \in \mathbb{C}^{N_r \times N_r}$, and $\mathbf{H}_{t,l} \in \mathbb{C}^{N_d \times N_r}$ denote the l th tap of the time-domain (TD) source precoding filter, the S - R channel, the TD relay precoding filter, and the R - D channel, respectively. The noise vectors at R and D are denoted by

$$\begin{aligned} \mathbf{v} &= [\mathbf{v}_0^T, \dots, \mathbf{v}_{N_c-1}^T]^T \sim \mathcal{CN}(\mathbf{0}, \sigma_v^2 \mathbf{I}_{N_r N_c}), \\ \mathbf{u} &= [\mathbf{u}_0^T, \dots, \mathbf{u}_{N_c-1}^T]^T \sim \mathcal{CN}(\mathbf{0}, \sigma_u^2 \mathbf{I}_{N_d N_c}), \end{aligned} \quad (2)$$

where $\mathbf{v}_n = [v_n(1), v_n(2), \dots, v_n(N_r)]^T$ and $\mathbf{u}_n = [u_n(1), u_n(2), \dots, u_n(N_d)]^T$ denote the additive white Gaussian noise (AWGN) vectors at R and D at time n , respectively. The block circular matrices $\{\mathbf{P}_t, \mathbf{G}_t, \mathbf{A}_t, \mathbf{H}_t\}$ can be decomposed as

$$\begin{aligned} \mathbf{P}_t &= \mathbf{F}_{N_s}^\dagger \mathbf{P}_f \mathbf{F}_M, & \mathbf{G}_t &= \mathbf{F}_{N_r}^\dagger \mathbf{G}_f \mathbf{F}_{N_s}, \\ \mathbf{A}_t &= \mathbf{F}_{N_r}^\dagger \mathbf{A}_f \mathbf{F}_{N_r}, & \mathbf{H}_t &= \mathbf{F}_{N_d}^\dagger \mathbf{H}_f \mathbf{F}_{N_r}, \end{aligned} \quad (3)$$

with $\mathbf{F}_\Upsilon = \mathbb{F}_{N_\Upsilon}^\dagger \otimes \mathbf{I}_\Upsilon$, $\Upsilon \in \{M, N_s, N_r, N_d\}$, $\mathbf{X}_f = \text{blkdiag}([\mathbf{X}_0^T, \dots, \mathbf{X}_{N_c-1}^T]^T)$, and $\mathbf{X}_f \in \{\mathbf{P}_f, \mathbf{G}_f, \mathbf{A}_f, \mathbf{H}_f\}$. Here, $\mathbf{P}_k \in \mathbb{C}^{N_s \times M}$, $\mathbf{G}_k \in \mathbb{C}^{N_r \times N_s}$, $\mathbf{A}_k \in \mathbb{C}^{N_r \times N_r}$, and $\mathbf{H}_k \in \mathbb{C}^{N_d \times N_r}$ represent the frequency-domain (FD) source precoding, S - R channel, relay precoding, and R - D channel matrices for the k th frequency tone, respectively. We define the equivalent end-to-end channel matrix $\mathbf{Q}_t = \mathbf{H}_t \mathbf{A}_t \mathbf{G}_t \mathbf{P}_t$ and express it as $\mathbf{Q}_t = \mathbf{F}_{N_d}^\dagger \mathbf{Q}_f \mathbf{F}_M$, where $\mathbf{Q}_f = \text{blkdiag}([\mathbf{Q}_0^T, \dots, \mathbf{Q}_{N_c-1}^T]^T)$ with $\mathbf{Q}_k = \mathbf{H}_k \mathbf{A}_k \mathbf{G}_k \mathbf{P}_k \in \mathbb{C}^{N_d \times M}$ representing the equivalent S - D channel matrix on the k th frequency tone. Furthermore, the CV matrix of the equivalent noise vector $\mathbf{n} = \mathbf{H}_t \mathbf{A}_t \mathbf{v} + \mathbf{u}$ can be obtained as

$$\mathbf{K} = E[\mathbf{n}\mathbf{n}^\dagger] = \mathbf{F}_{N_d}^\dagger \mathbf{K}_f \mathbf{F}_{N_d}, \quad (4)$$

where $\mathbf{K}_f = \sigma_v^2 \mathbf{H}_f \mathbf{A}_f \mathbf{A}_f^\dagger \mathbf{H}_f^\dagger + \sigma_u^2 \mathbf{I}_{N_d N_c}$.

B. Equalization at the Destination

The received signal \mathbf{y} is transformed into the FD using \mathbf{F}_{N_d} and equalized by an FD feedforward filter (FFF) $\mathbf{W}_f = \text{blkdiag}([\mathbf{W}_0^T, \dots, \mathbf{W}_{N_c-1}^T]^T)$. The resulting signal is then transformed into the TD using \mathbf{F}_M^\dagger resulting in

$$\hat{\mathbf{y}} = \mathbf{W}_t \mathbf{y}, \quad (5)$$

where $\mathbf{W}_t = \mathbf{F}_M^\dagger \mathbf{W}_f \mathbf{F}_{N_d}$ is the equivalent TD FFF and $\hat{\mathbf{y}} = [\hat{\mathbf{y}}_0^T, \dots, \hat{\mathbf{y}}_{N_c-1}^T]^T$ with $\hat{\mathbf{y}}_n = [\hat{y}_n(1), \hat{y}_n(2), \dots, \hat{y}_n(M)]^T$ denoting the n th signal vector at the output of the FFF. If FD-LE is employed, $\hat{\mathbf{y}}_n$ is the decision variable for the n th source symbol vector. On the

¹For simplicity of presentation, the CP insertion is not shown in Fig. 1.

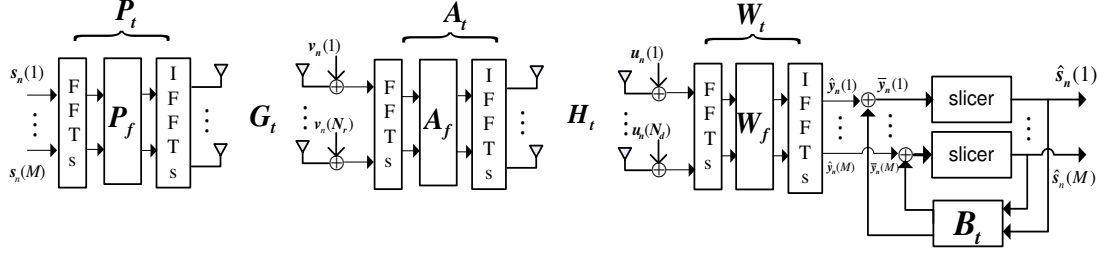


Fig. 1. System model for a MIMO relay system with SC-FDE at the destination.

other hand, for FD-DFE, \hat{y}_n is further processed using a TD feedback filter (FBF) to perform interference cancellation. Assuming correct feedback at the output of the slicer², the signal corresponding to the m th data stream at time n at the input of the slicer is given by

$$\bar{y}_n(m) = \hat{y}_n(m) - \sum_{l=0}^{N_{fb}} [\mathbf{B}_{t,l}]_{(m,:)} \mathbf{s}_{(n-l) \bmod N_c}, \quad (6)$$

where $\mathbf{B}_{t,l}$ denotes the coefficient matrix of the l th tap of the FBF, $[\mathbf{X}]_{(m,:)}$ stands for the m th row of matrix \mathbf{X} , N_{fb} is the number of feedback taps, and $(\cdot) \bmod N$ denotes the modulo- N operation. From (6) we observe that at the initial stage of the feedback process, i.e., when $n = 0$, $[\mathbf{s}_{N_c - N_{fb}}, \dots, \mathbf{s}_{N_c}]$ has to be known *a priori*, which can be accomplished by using known training symbols. Nevertheless, for detection of $s_n(m)$, $[s_n(1), \dots, s_n(m-1)]$ is still unknown. Therefore, for causal detection, the 0th tap of the FBF, i.e., $\mathbf{B}_{t,0}$, has to be a lower triangular matrix with zero diagonal entries. By collecting all $\bar{y}_n(m)$ into a vector $\bar{\mathbf{y}} = [\bar{\mathbf{y}}_0^T, \dots, \bar{\mathbf{y}}_{N_c-1}^T]^T$ with $\bar{\mathbf{y}}_n = [\bar{y}_n(1), \bar{y}_n(2), \dots, \bar{y}_n(M)]^T$, we arrive at

$$\bar{\mathbf{y}} = \hat{\mathbf{y}} - \mathbf{B}_t \mathbf{s}, \quad (7)$$

where $\mathbf{B}_t = \text{blkcirc}([\mathbf{B}_{t,0}^T, \dots, \mathbf{B}_{t,N_{fb}}^T, \mathbf{0}_{M \times M(N_c - N_{fb} - 1)}]^T) \in \mathbb{C}^{MN_c \times MN_c}$ is the equivalent TD-FBF. Thus, the error vector at the input of the slicer can be expressed as

$$\mathbf{e} = \bar{\mathbf{y}} - \mathbf{s} = \hat{\mathbf{y}} - \underbrace{(\mathbf{B}_t + \mathbf{I}_{MN_c})}_{\mathbf{C}_t} \mathbf{s} = \hat{\mathbf{y}} - \mathbf{C}_t \mathbf{s}, \quad (8)$$

where $\mathbf{C}_t = \text{blkcirc}([\mathbf{C}_{t,0}^T, \dots, \mathbf{C}_{t,N_{fb}}^T, \mathbf{0}^T]^T)$ with $\mathbf{C}_{t,n} = \mathbf{B}_{t,n}, \forall n \neq 0$ and $\mathbf{C}_{t,0} = \mathbf{B}_{t,0} + \mathbf{I}_M$. The block circular matrix \mathbf{C}_t can be decomposed as $\mathbf{C}_t = \mathbf{F}_M^\dagger \mathbf{C}_f \mathbf{F}_M$, where $\mathbf{C}_f = \text{blkdiag}([\mathbf{C}_0^T, \dots, \mathbf{C}_{N_c-1}^T]^T)$. We note that by setting $\mathbf{C}_t = \mathbf{I}_{MN_c}$, FD-DFE reduces to FD-LE.

III. OPTIMAL MINIMUM MSE FDE FILTER DESIGN

In this section, we derive the optimal minimum MSE equalization filters at the destination and the corresponding error CV matrices at the output of the equalizer as functions of the source and relay precoding matrices. Combining (1)-(5) and (8), the error CV matrix, $\mathbf{E} \triangleq E[\mathbf{e}\mathbf{e}^\dagger]$, can be expressed

as

$$\mathbf{E} = \mathbf{F}_M^\dagger \left(\mathbf{W}_f (\sigma_s^2 \mathbf{Q}_f \mathbf{Q}_f^\dagger + \mathbf{K}_f) \mathbf{W}_f^\dagger - \sigma_s^2 \mathbf{W}_f \mathbf{Q}_f \mathbf{C}_f^\dagger - \sigma_s^2 \mathbf{C}_f \mathbf{Q}_f^\dagger \mathbf{W}_f^\dagger + \sigma_s^2 \mathbf{C}_f \mathbf{C}_f^\dagger \right) \mathbf{F}_M. \quad (9)$$

Following the conventional equalization design methodology, the optimum FD FFF is obtained by minimizing the sum of stream MSEs, $\text{tr}(\mathbf{E})$, which yields

$$\mathbf{W}_f^* = \sigma_s^2 \mathbf{C}_f \mathbf{Q}_f^\dagger \left(\sigma_s^2 \mathbf{Q}_f \mathbf{Q}_f^\dagger + \mathbf{K}_f \right)^{-1}. \quad (10)$$

Substituting \mathbf{W}_f^* into (9) and simplifying the resulting expression, the CV matrix can be rewritten as

$$\mathbf{E} = \sigma_s^2 \mathbf{F}_M^\dagger \mathbf{C}_f \boldsymbol{\Psi}_f^{-1} \mathbf{C}_f^\dagger \mathbf{F}_M, \quad (11)$$

where $\boldsymbol{\Psi}_f = \text{blkdiag}([\boldsymbol{\Psi}_0^T, \dots, \boldsymbol{\Psi}_{N_c-1}^T]^T) \in \mathbb{C}^{MN_c \times MN_c}$ with

$$\boldsymbol{\Psi}_k = \sigma_s^2 \mathbf{Q}_k^\dagger \left(\sigma_v^2 \mathbf{H}_k \mathbf{A}_k \mathbf{A}_k^\dagger \mathbf{H}_k^\dagger + \sigma_u^2 \mathbf{I}_{N_d} \right)^{-1} \mathbf{Q}_k + \mathbf{I}_M. \quad (12)$$

From (11) we observe that \mathbf{E} is a block circular matrix. Hence, its block diagonal entries, $\mathbf{E}_n \in \mathbb{C}^{M \times M}, \forall n$, are identical, i.e., $\mathbf{E}_n = \hat{\mathbf{E}}, \forall n$. Since the diagonal entries of \mathbf{E}_n represent the MSEs of the different spatial streams at time n , symbols from the same stream experience identical MSEs. CV matrix $\hat{\mathbf{E}}$ can be conveniently written as

$$\hat{\mathbf{E}} = \frac{\sigma_s^2}{N_c} \sum_{k=1}^{N_c} \mathbf{C}_k^\dagger \boldsymbol{\Psi}_k^{-1} \mathbf{C}_k. \quad (13)$$

A. CV Matrix and Filter Design for FD-LE

Eqs. (10) and (13) are valid for both FD-LE and FD-DFE. For the special case of FD-LE, we can set $\mathbf{C}_f = \mathbf{I}_{MN_c}$, which leads to

$$\mathbf{W}_f^* = \sigma_s^2 \mathbf{Q}_f^\dagger \left(\sigma_s^2 \mathbf{Q}_f \mathbf{Q}_f^\dagger + \mathbf{K}_f \right)^{-1} \quad (14)$$

and CV matrix

$$\hat{\mathbf{E}}_{\text{FD-LE}} = \frac{\sigma_s^2}{N_c} \sum_{k=1}^{N_c} \boldsymbol{\Psi}_k^{-1}. \quad (15)$$

Interestingly, $\hat{\mathbf{E}}_{\text{FD-LE}}$ is equal to the arithmetic mean of the subcarrier CV matrices, $\boldsymbol{\Psi}_k^{-1}$, in MIMO-OFDM relay systems [6].

²Correct feedback is a common assumption for the design of decision feedback equalizers [11]-[13].

B. CV Matrix and Filter Design for FD-DFE

The FD-DFE CV matrix depends on the FD-FBF matrices \mathbf{C}_k . Since the FBF has to be implemented in the TD, we express \mathbf{C}_k in terms of the TD FBF coefficients $\mathbf{C}_{t,n}$ as $\mathbf{C}_k = \sum_{n=0}^{N_{fb}} \mathbf{C}_{t,n} e^{-j\frac{2\pi}{N_c}nk}$. Now, (13) can be rewritten as

$$\begin{aligned} \hat{\mathbf{E}}_{\text{FD-DFE}} &= \frac{\sigma_s^2}{N_c} \sum_{k=0}^{N_c-1} \left[\sum_{n=0}^{N_{fb}} \mathbf{C}_{t,n} e^{-j\frac{2\pi}{N_c}nk} \boldsymbol{\Psi}_k^{-1} \sum_{m=0}^{N_{fb}} \mathbf{C}_{t,m}^\dagger e^{-j\frac{2\pi}{N_c}mk} \right] \\ &= \frac{\sigma_s^2}{N_c} \sum_{n=0}^{N_{fb}} \sum_{m=0}^{N_{fb}} \left(\mathbf{C}_{t,n} \sum_{k=0}^{N_c-1} \boldsymbol{\Psi}_k^{-1} e^{-j\frac{2\pi}{N_c}(n-m)k} \mathbf{C}_{t,m}^\dagger \right) \\ &= \frac{\sigma_s^2}{N_c} \hat{\mathbf{C}} \mathbf{Z} \hat{\mathbf{C}}^\dagger. \end{aligned} \quad (16)$$

To simplify the notation, we have used the definitions $\hat{\mathbf{C}} = [\mathbf{C}_{t,0}, \dots, \mathbf{C}_{t,N_{fb}}]$ and

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}_0 & \mathbf{z}_1 & \dots & \mathbf{z}_{N_{fb}} \\ \mathbf{z}_1^\dagger & \mathbf{z}_0 & \dots & \mathbf{z}_{N_{fb}-1} \\ \vdots & & \ddots & \vdots \\ \mathbf{z}_{N_{fb}}^\dagger & \mathbf{z}_{N_{fb}-1}^\dagger & \dots & \mathbf{z}_0 \end{bmatrix}, \quad (17)$$

where $\mathbf{z}_n = \sum_{k=0}^{N_c-1} \boldsymbol{\Psi}_k^{-1} e^{j\frac{2\pi}{N_c}kn}$. The optimal $\hat{\mathbf{C}}$ minimizing $\text{tr}\{\mathbf{E}_{\text{FD-DFE}}\}$ can be obtained by solving

$$\min_{\hat{\mathbf{C}} \boldsymbol{\Theta} = \mathbf{C}_{t,0}} \text{tr}(\hat{\mathbf{C}} \mathbf{Z} \hat{\mathbf{C}}^\dagger), \quad (18)$$

where $\boldsymbol{\Theta} = [\mathbf{I}_M, \mathbf{0}_{M \times (N_{fb}-1)}]$. Problem (18) can be solved using the standard Lagrange multiplier method, leading to [13]

$$\hat{\mathbf{C}}^* = \mathbf{C}_{t,0} (\boldsymbol{\Theta} \mathbf{Z}^{-1} \boldsymbol{\Theta})^{-1} \boldsymbol{\Theta}^\dagger \mathbf{Z}^{-1}. \quad (19)$$

By partitioning \mathbf{Z} and \mathbf{Z}^{-1} as

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_{11} & \mathbf{Z}_{12} \\ \mathbf{Z}_{12}^\dagger & \mathbf{Z}_{22} \end{bmatrix}, \mathbf{Z}^{-1} = \begin{bmatrix} \mathbf{U}_{11} & \mathbf{U}_{12} \\ \mathbf{U}_{12}^\dagger & \mathbf{U}_{22} \end{bmatrix}, \quad (20)$$

where $\{\mathbf{Z}_{11}, \mathbf{U}_{11}\} \in \mathbb{C}^{M \times M}$, $\{\mathbf{Z}_{12}, \mathbf{U}_{12}\} \in \mathbb{C}^{M \times MN_{fb}}$, and $\{\mathbf{Z}_{22}, \mathbf{U}_{22}\} \in \mathbb{C}^{MN_{fb} \times MN_{fb}}$, $\hat{\mathbf{C}}$ can be further expressed as

$$\hat{\mathbf{C}} = [\mathbf{C}_{t,0}^{-1}, \quad \mathbf{C}_{t,0}^{-1} \mathbf{Z}_{12} \mathbf{Z}_{22}^{-1}]. \quad (21)$$

Substituting (21) into (16), the FD-DFE CV matrix can be rewritten as

$$\mathbf{E}_{\text{FD-DFE}} = \frac{\sigma_s^2}{N_c} \mathbf{C}_{t,0} \mathbf{U}_{11}^{-1} \mathbf{C}_{t,0}^\dagger. \quad (22)$$

To complete the FBF design, the optimal $\mathbf{C}_{t,0}$ has to be determined. To this end, we introduce the Cholesky decomposition of \mathbf{U}_{11}^{-1} as

$$\mathbf{U}_{11}^{-1} = \mathbf{L} \mathbf{D} \mathbf{L}^\dagger, \quad (23)$$

where \mathbf{L} is a unit-diagonal lower triangular matrix and \mathbf{D} is a diagonal matrix with positive main diagonal entries. Now, it is easy to verify that the optimal $\mathbf{C}_{t,0}$ which minimizes $\text{tr}(\mathbf{E}_{\text{FD-DFE}})$ is given by $\mathbf{C}_{t,0}^* = \mathbf{L}^{-1}$. Hence, the optimal $\hat{\mathbf{C}}$ is obtained as

$$\hat{\mathbf{C}}^* = [\mathbf{L}^{-1}, \quad \mathbf{L}^{-1} \mathbf{Z}_{12} \mathbf{Z}_{22}^{-1}]. \quad (24)$$

The structure of the optimal FBF can be interpreted as follows: $\mathbf{L}^{-1} \in \mathbb{C}^{M \times M}$ is a lower-triangular matrix which cancels the inter-antenna interference (IAI) in the current time slot, and the remaining FBF coefficients, $\mathbf{L}^{-1} \mathbf{Z}_{12} \mathbf{Z}_{22}^{-1} \in \mathbb{C}^{M \times MN_{fb}}$, cancel both the IAI and inter-symbol interference (ISI) stemming from the previous $N_{fb} - 1$ time slots. Inserting $\hat{\mathbf{C}}^*$ into (22) the error CV matrix can be written as

$$\hat{\mathbf{E}}_{\text{FD-DFE}} = \frac{\sigma_s^2}{N_c} \mathbf{D}. \quad (25)$$

Interestingly, unlike for FD-LE, the error CV matrix for FD-DFE is a diagonal matrix, which also depends on the number of FBF taps N_{fb} .

IV. SOURCE AND RELAY PRECODING MATRIX OPTIMIZATION

Exploiting the expressions for the error CV matrix obtained in the previous section, in this section, we minimize a general function $f(\text{diag}[\hat{\mathbf{E}}])$ of the spatial stream MSEs at the output of the equalization filter under separate constraints on the powers consumed at the source and the relay³. Mathematically, the optimization problem is stated as

$$\begin{aligned} \min_{\{\mathbf{P}_k, \mathbf{A}_k\}} \quad & f(\text{diag}[\hat{\mathbf{E}}]) \\ \text{s.t.} \quad & \text{tr}(\mathbf{E}[\mathbf{x}\mathbf{x}^\dagger]) \leq P_S, \quad \text{tr}(\mathbf{E}[\mathbf{t}\mathbf{t}^\dagger]) \leq P_R, \end{aligned} \quad (26)$$

where $\hat{\mathbf{E}} = \hat{\mathbf{E}}_{\text{FD-LE}}$ and $\hat{\mathbf{E}} = \hat{\mathbf{E}}_{\text{FD-DFE}}$ for FD-LE and FD-DFE, respectively, P_S and P_R are the power budgets for S and R , respectively, and $\text{diag}[\mathbf{M}]$ denotes a vector containing the main diagonal entries of matrix \mathbf{M} . The objective function $f(\text{diag}[\hat{\mathbf{E}}])$ can be either a Schur-convex or a Schur-concave increasing function with respect to (w.r.t.) $\text{diag}[\hat{\mathbf{E}}]$ [14]. For concreteness, in this paper, we consider the three most important objective functions of this type, namely the arithmetic MSE (AMSE), the geometric MSE (GMSE), and the maximum MSE (maxMSE)

$$f(\text{diag}[\hat{\mathbf{E}}]) = \begin{cases} \sum_{m=1}^M \hat{\mathbf{E}}_{mm}, & \text{AMSE} \\ \prod_{m=1}^M \hat{\mathbf{E}}_{mm}, & \text{GMSE} \\ \max_{m=1}^M \hat{\mathbf{E}}_{mm}, & \text{maxMSE} \end{cases}, \quad (27)$$

where $\hat{\mathbf{E}}_{mm}$ denotes the m th diagonal entry of $\hat{\mathbf{E}}$. The AMSE and GMSE are Schur-concave functions while the maxMSE is a Schur-convex function w.r.t. $\text{diag}[\hat{\mathbf{E}}]$ [14]. We note that similar objective functions have been considered for MIMO-OFDM based relay systems in [6]. However, for MIMO-OFDM based relay systems, the AMSE, GMSE, and maxMSE are the sum, product, and maximum of the subcarrier MSEs of different spatial sub-streams. In contrast, in (27), these three quantities are the the sum, product, and maximum of the sub-stream MSEs of a single carrier.

³We note that our derivations can be extended to a joint source and relay power constraint. While such a joint power constraint offers more degrees of freedom for the system design, separate power constraints appear more practical since usually the source node and the relay node have their own power supplies.

The power consumptions at source and relay are given by

$$\begin{aligned} \text{tr}(E[\mathbf{x}\mathbf{x}^\dagger]) &= \sigma_s^2 \sum_{k=0}^{N_c-1} \text{tr}(\mathbf{P}_k \mathbf{P}_k^\dagger), \\ \text{tr}(E[\mathbf{t}\mathbf{t}^\dagger]) &= \sum_{k=0}^{N_c-1} \text{tr}(\mathbf{A}_k (\sigma_s^2 \mathbf{G}_k \mathbf{P}_k \mathbf{P}_k^\dagger \mathbf{G}_k^\dagger + \sigma_v^2 \mathbf{I}_{N_r}) \mathbf{A}_k^\dagger). \end{aligned} \quad (28)$$

Since the optimization variables in (26) are matrices, solving the problem directly would incur high complexity. In the following, we will first derive the structure of the optimal precoding matrices. Knowing this structure will allow us to transform the optimization problem into an optimization problem with scalar variables.

A. Structure of the Optimal Precoding Matrices for FD-LE

We first derive the structure of the optimal source and relay precoding matrices for FD-LE. We begin by introducing the following singular-value decompositions (SVDs) of the channel matrices

$$\mathbf{G}_k = \mathbf{U}_G^{(k)} \mathbf{\Lambda}_G^{(k)} \mathbf{V}_G^{(k)\dagger}, \quad \mathbf{H}_k = \mathbf{U}_H^{(k)} \mathbf{\Lambda}_H^{(k)} \mathbf{V}_H^{(k)\dagger}, \quad \forall k, \quad (29)$$

where $\mathbf{U}_G^{(k)} \in \mathbb{C}^{N_r \times N_r}$, $\mathbf{V}_G^{(k)} \in \mathbb{C}^{N_s \times N_s}$ and $\mathbf{U}_H^{(k)} \in \mathbb{C}^{N_d \times N_d}$, $\mathbf{V}_H^{(k)} \in \mathbb{C}^{N_r \times N_r}$ are the singular-vector matrices of \mathbf{G}_k and \mathbf{H}_k , respectively. Furthermore, $\mathbf{\Lambda}_G^{(k)} \in \mathbb{C}^{N_r \times N_s}$ and $\mathbf{\Lambda}_H^{(k)} \in \mathbb{C}^{N_d \times N_r}$ are the singular-value matrices of \mathbf{G}_k and \mathbf{H}_k , respectively, and have both increasing main diagonal elements.

Theorem 1: For the optimization problem in (26), the following structures of \mathbf{P}_k and \mathbf{A}_k are optimal

$$\mathbf{P}_k^* = \bar{\mathbf{V}}_G^{(k)} \mathbf{\Lambda}_P^{(k)} \mathbf{V}_0, \quad \mathbf{A}_k^* = \bar{\mathbf{V}}_H^{(k)} \mathbf{\Lambda}_A^{(k)} \bar{\mathbf{U}}_G^{(k)}, \quad \forall k, \quad (30)$$

where $\bar{\mathbf{V}}_G^{(k)}$, $\bar{\mathbf{U}}_G^{(k)}$, and $\bar{\mathbf{V}}_H^{(k)}$ contain the M right-most columns of $\mathbf{V}_G^{(k)}$, $\mathbf{U}_G^{(k)}$, and $\mathbf{V}_H^{(k)}$, respectively. $\mathbf{\Lambda}_P^{(k)}$ and $\mathbf{\Lambda}_A^{(k)}$ are $M \times M$ diagonal matrices with the m th diagonal element denoted by p_{km} and a_{km} , respectively. For Schur-concave functions, $\mathbf{V}_0 = \mathbf{I}_M$. For Schur-convex functions, \mathbf{V}_0 is a unitary matrix chosen in such a way that all main diagonal entries of $\hat{\mathbf{E}}$ are equal⁴.

Proof: Please refer to the Appendix. ■

Theorem 1 implies that for Schur-concave functions, the source and relay precoding matrices jointly diagonalize the MIMO relay channels at each frequency tone, while for Schur-convex functions, the precoding matrices diagonalize the channels up to a unitary rotation at the source. Therefore, the original optimization problem involving matrix variables can be transformed into a scalar power optimization problem across different spatial beams and frequency tones.

B. Transformation of Optimization Problem for FD-LE

Since the maxMSE is a Schur-convex function, according to Theorem 1, the unitary matrix, \mathbf{V}_0 , should be chosen to make all diagonal entries of $\hat{\mathbf{E}}$ equal. Recall from Section III that

$\text{diag}[\hat{\mathbf{E}}]$ represents the MSE of different spatial streams and all symbols of a particular stream have the same MSE. This means that for maxMSE, identical MSE is achieved for all symbols in the SC-FDE system. Hence, the remaining maxMSE power allocation problem is identical to that for the AMSE criterion. The only difference between the solutions for maxMSE and AMSE minimization lies in the choice of \mathbf{V}_0 . We note that this is not true for MIMO-OFDM relay systems, where the unitary transformation at the source only achieves identical spatial MSEs on each subcarrier, while the MSEs across subcarriers are in general different. To balance these MSEs, multilevel waterfilling has to be carried out in such MIMO-OFDM relay systems, which entails a much higher complexity compared to the single-level waterfilling required for the AMSE criterion, cf. [6]. Additionally, for MIMO-OFDM relay systems, the unitary rotation matrices are in general different on each subcarrier as the number of transmitted data streams may vary from subcarrier to subcarrier. However, for SC-FDE, the rotation matrices are identical for all frequency tones since the number of data streams is determined in the time domain.

Because of the equivalence of the power allocation problems for maxMSE and AMSE, in the following, we focus on the power allocation problem for the AMSE and GMSE criteria. From (12) and (30), we obtain $\Psi_k = \mathbf{V}_0^\dagger \Phi_k \mathbf{V}_0$ with

$$\Phi_k = \sigma_s^2 \mathbf{\Lambda}_P^{(k)2} \bar{\mathbf{\Lambda}}_G^{(k)2} \mathbf{\Lambda}_A^{(k)2} \bar{\mathbf{\Lambda}}_H^{(k)2} \left(\sigma_v^2 \mathbf{\Lambda}_A^{(k)2} \bar{\mathbf{\Lambda}}_H^{(k)2} + \sigma_u^2 \mathbf{I}_M \right)^{-1} + \mathbf{I}_M, \quad (31)$$

where $\bar{\mathbf{\Lambda}}_G^{(k)}$ and $\bar{\mathbf{\Lambda}}_H^{(k)}$ are diagonal matrices whose diagonal entries contain the M largest singular values of $\mathbf{G}^{(k)}$ and $\mathbf{H}^{(k)}$, respectively. Now, we can rewrite the objective functions as

$$f_X(\Phi) = \begin{cases} \sum_{m=1}^M \left(\frac{1}{N_c} \sum_{k=0}^{N_c-1} \Phi_{km}^{-1} \right), & \text{X=AMSE} \\ \sum_{m=1}^M \log_2 \left(\frac{1}{N_c} \sum_{k=0}^{N_c-1} \Phi_{km}^{-1} \right), & \text{X=GMSE} \end{cases}, \quad (32)$$

where $\Phi = \{\Phi_{km}, \forall k, m\}$ with

$$\Phi_{km} = \frac{\sigma_s^2 p_{km}^2 g_{km}^2 a_{km}^2 h_{km}^2}{\sigma_v^2 a_{km}^2 h_{km}^2 + \sigma_u^2} + 1. \quad (33)$$

Here, g_{km} and h_{km} denote the m th main diagonal elements of $\bar{\mathbf{\Lambda}}_G^{(k)}$ and $\bar{\mathbf{\Lambda}}_H^{(k)}$, respectively, and represent the corresponding channel gains of the m th spatial stream on the k th frequency tone. Note that, for the GMSE criterion, we have taken the logarithm of the original objective function to facilitate the subsequent optimization. Due to the monotonicity of the logarithm, the new objective function has the same optimal solution as the original one. The new objective function can be rewritten as

$$- \sum_{m=1}^M \log_2 (\text{SINR}_m + 1), \quad (34)$$

where $\text{SINR}_m = \left(\frac{1}{N_c} \sum_{k=0}^{N_c-1} \Phi_{km}^{-1} \right)^{-1} - 1$ is the signal-to-interference-plus-noise ratio (SINR) of the m th data stream. This implies that (34) is essentially the negative sum of the channel capacities of different spatial streams. Therefore, minimization of the GMSE is equivalent to the maximization

⁴ In practice, \mathbf{V}_0 can be chosen as a DFT matrix or a Hadamard matrix with appropriate dimensions.

of the capacity of the considered MIMO SC-FDE relay system. By exploiting (30), the expression for the power consumption on the left hand side of the constraints in (26) can be expressed as

$$\begin{aligned} \text{tr}(E[\mathbf{x}\mathbf{x}^\dagger]) &= \sigma_s^2 \sum_{k=0}^{N_c-1} \text{tr}(\mathbf{\Lambda}_P^{(k)2}) = \sum_{k=0}^{N_c-1} \sum_{m=1}^M P_{s,km} \\ \text{tr}(E[\mathbf{t}\mathbf{t}^\dagger]) &= \sum_{k=0}^{N_c-1} \text{tr}(\mathbf{\Lambda}_A^{(k)2} (\sigma_s^2 \mathbf{\Lambda}_P^{(k)2} \bar{\mathbf{\Lambda}}_G^{(k)2} + \sigma_u^2 \mathbf{I}_M)) \\ &= \sum_{k=0}^{N_c-1} \sum_{m=1}^M P_{r,km}, \end{aligned} \quad (35)$$

where

$$P_{s,km} = \sigma_s^2 p_{km}^2, \quad P_{r,km} = a_{km}^2 (\sigma_s^2 p_{km}^2 g_{km}^2 + \sigma_v^2) \quad (36)$$

can be interpreted as the power allocated to the k th frequency tone and the m th spatial stream at the source and the relay, respectively. By rewriting Φ_{km} in terms of the newly introduced variables $P_{s,km}$ and $P_{r,km}$ as

$$\Phi_{km} = \frac{P_{s,km} P_{r,km} g_{km}^2 h_{km}^2}{\sigma_v^2 P_{r,km} h_{km}^2 + \sigma_u^2 (P_{s,km} g_{km}^2 + \sigma_v^2)} + 1, \quad (37)$$

problem (26) can be reformulated as the following power allocation problem

$$\begin{aligned} \min_{\{P_{s,km}, P_{r,km}\}} \quad & f_X(\Phi) \\ \text{s.t.} \quad & \sum_{k=0}^{N_c-1} \sum_{m=1}^M P_{s,km} \leq P_S, \\ & \sum_{k=0}^{N_c-1} \sum_{m=1}^M P_{r,km} \leq P_R, \\ & P_{s,km} \geq 0, P_{r,km} \geq 0, \forall k, m, \end{aligned} \quad (38)$$

where the constraints $P_{s,km} \geq 0, P_{r,km} \geq 0, \forall k, m$, ensure that the allocated powers are not negative.

C. Structure of the Optimal Precoding Matrices for FD-DFE

For the FD-DFE receiver, we observe from (25) that $\mathbf{E}_{\text{FD-DFE}}$ is not an explicit function of optimization variables \mathbf{P}_k and \mathbf{A}_k , which renders the optimization a challenging task. In the following, we will show that by using some proper transformations, an upper bound for the original objective function can be derived, which is equivalent to one of the objective functions considered for the FD-LE receiver. However, first we will show that for FD-DFE, the three considered objective functions are equivalent.

1) *Equivalence of Objective Functions:* Since $\mathbf{E}_{\text{FD-DFE}}$ in (25) is a diagonal matrix, we invoke the following matrix arithmetic-geometric mean inequality

$$\frac{1}{M} \text{tr}(\mathbf{D}) = \frac{1}{M} \sum_{i=1}^M [\mathbf{D}]_{(i,i)} \geq \left(\prod_{i=1}^M [\mathbf{D}]_{(i,i)} \right)^{\frac{1}{M}} = \det(\mathbf{D})^{\frac{1}{M}}, \quad (39)$$

where equality holds if and only if (i.f.f.) all main diagonal elements of \mathbf{D} are equal. The inequality provides some important

insights into the objective function for FD-DFE. First, it implies that the AMSE, i.e., $\text{tr}(\mathbf{D})$, is lower bounded by the term involving the GMSE, i.e., $\det(\mathbf{D})$. Second, this lower bound is achieved i.f.f. the MSEs of all streams are identical. Therefore, making the diagonal entries of \mathbf{D} identical will enable us to minimize the AMSE, GMSE, and maxMSE simultaneously. Consequently, for FD-DFE, the three considered objective functions become equivalent if this condition is fulfilled, and the value of the objective function is equal to $\det(\mathbf{D})^{\frac{1}{M}}$. In the sequel, we will show how this can be achieved by properly choosing a unitary matrix at the source precoder. By noting that

$$\det(\mathbf{D}) = \det(\mathbf{L}\mathbf{D}\mathbf{L}^\dagger) = \det(\mathbf{U}_{11}^{-1}), \quad (40)$$

where we have exploited the properties that $\det(\mathbf{A}\mathbf{B}) = \det(\mathbf{A})\det(\mathbf{B})$ and $\det(\mathbf{L}) = 1$ [18], we can obtain

$$\mathbf{U}_{11}^{-1} = \mathbf{L}\mathbf{D}^{1/2}(\mathbf{L}\mathbf{D}^{1/2})^\dagger = (\mathbf{Q}\mathbf{R})^\dagger \mathbf{Q}\mathbf{R}, \quad (41)$$

where \mathbf{Q} is an arbitrary unitary matrix of appropriate dimension and $\mathbf{R} = (\mathbf{L}\mathbf{D}^{1/2})^\dagger$ is a lower triangular matrix whose main diagonal elements are equal to the square root of the main diagonal elements of \mathbf{D} . Therefore, finding a matrix \mathbf{D} with equal diagonal elements is equivalent to finding a matrix \mathbf{R} with equal diagonal elements. Since it follows from (41) that $\mathbf{U}_{11}^{-1/2} = \mathbf{Q}\mathbf{R}$, we have to find a matrix decomposition of $\mathbf{U}_{11}^{-1/2}$ such that \mathbf{R} has identical diagonal elements. Such a decomposition is referred to as equal-diagonal QR decomposition (E-QRD) [16] or geometric-mean decomposition (GMD) [17]. Note that for an arbitrary matrix $\mathbf{\Xi}$, the standard form of the GMD is,

$$\mathbf{\Xi}\mathbf{V}_1^\dagger = \mathbf{Q}\mathbf{R}, \quad (42)$$

where \mathbf{V}_1 is a unitary matrix chosen to make the diagonal entries of \mathbf{R} all equal. Therefore, to facilitate the application of GMD, we rewrite $\mathbf{U}_{11}^{-1/2}$ as

$$\mathbf{U}_{11}^{-1/2} = \mathbf{\Xi}\mathbf{V}_1^\dagger \quad (43)$$

In the following, we will find the explicit expression for $\mathbf{\Xi}$ to perform the decomposition in (42). By expressing \mathbf{P}_k as the product of \mathbf{V}_1 and a general matrix $\tilde{\mathbf{P}}_k$,

$$\mathbf{P}_k = \tilde{\mathbf{P}}_k \mathbf{V}_1^\dagger, \quad (44)$$

we can write Ψ_k as $\Psi_k = \mathbf{V}_1 \hat{\Psi}_k \mathbf{V}_1^\dagger$, where

$$\begin{aligned} \hat{\Psi}_k &= \sigma_s^2 \tilde{\mathbf{P}}_k^\dagger \mathbf{G}_k^\dagger \mathbf{A}_k^\dagger \mathbf{H}_k^\dagger \left(\sigma_v^2 \mathbf{H}_k \mathbf{A}_k \mathbf{A}_k^\dagger \mathbf{H}_k^\dagger + \sigma_u^2 \mathbf{I}_{N_d} \right)^{-1} \\ &\quad \mathbf{H}_k \mathbf{A}_k \mathbf{G}_k \tilde{\mathbf{P}}_k + \mathbf{I}_M. \end{aligned} \quad (45)$$

Note that $\hat{\Psi}_k$ has the same form of Ψ_k in (12) but with \mathbf{P}_k replaced by $\tilde{\mathbf{P}}_k$. Therefore, matrix \mathbf{Z} in (17) can be written as

$$\begin{aligned} \mathbf{Z} &= \begin{bmatrix} \mathbf{V}_1 \bar{\mathbf{z}}_0 \mathbf{V}_1^\dagger & \mathbf{V}_1 \bar{\mathbf{z}}_1 \mathbf{V}_1^\dagger & \cdots & \mathbf{V}_1 \bar{\mathbf{z}}_{N_{fb}} \mathbf{V}_1^\dagger \\ \mathbf{V}_1 \bar{\mathbf{z}}_1^\dagger \mathbf{V}_1^\dagger & \mathbf{V}_1 \bar{\mathbf{z}}_0 \mathbf{V}_1^\dagger & \cdots & \mathbf{V}_1 \bar{\mathbf{z}}_{N_{fb}-1} \mathbf{V}_1^\dagger \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{V}_1 \bar{\mathbf{z}}_{N_{fb}}^\dagger \mathbf{V}_1^\dagger & \mathbf{V}_1 \bar{\mathbf{z}}_{N_{fb}-1}^\dagger \mathbf{V}_1^\dagger & \cdots & \mathbf{V}_1 \bar{\mathbf{z}}_0 \mathbf{V}_1^\dagger \end{bmatrix} \\ &= (\mathbf{I}_{N_{fb}} \otimes \mathbf{V}_1) \bar{\mathbf{Z}} (\mathbf{I}_{N_{fb}} \otimes \mathbf{V}_1^\dagger), \end{aligned} \quad (46)$$

where $\bar{\mathbf{Z}}$ has the same form as \mathbf{Z} in (17) with \mathbf{z}_n replaced by $\bar{\mathbf{z}}_n = \sum_{k=0}^{N_c-1} \hat{\Psi}_k^{-1} e^{j\frac{2\pi}{N_c}kn}$. By noting that

$$\mathbf{Z}^{-1} = (\mathbf{I}_{N_{fb}} \otimes \mathbf{V}_1) \bar{\mathbf{Z}}^{-1} (\mathbf{I}_{N_{fb}} \otimes \mathbf{V}_1^\dagger), \quad (47)$$

where we have used $(\mathbf{I}_{N_{fb}} \otimes \mathbf{V}_1)^{-1} = \mathbf{I}_{N_{fb}} \otimes \mathbf{V}_1^\dagger$, we obtain from (20)

$$\mathbf{U}_{11}^{-1} = \mathbf{V}_1 \bar{\mathbf{U}}_{11}^{-1} \mathbf{V}_1^\dagger, \quad (48)$$

where $\bar{\mathbf{U}}_{11}$ is the first $M \times M$ submatrix of $\bar{\mathbf{Z}}^{-1}$. Comparing (48) with (43), we observe that $\bar{\mathbf{U}}_{11}^{-1/2}$ is the explicit form for Ξ . Hence, for a given $\bar{\mathbf{U}}_{11}$, we can always find a unitary matrix \mathbf{V}_1 which achieves the MSE lower bound.

As $\bar{\mathbf{U}}_{11}$ is a function of relay precoding matrix \mathbf{A}_k as well as the remaining part of the source precoding matrix, i.e., $\tilde{\mathbf{P}}_k$, in the following, we need to determine these matrices. To this end, we write $\mathbf{U}_{11} = (\mathbf{Z}_{11} - \mathbf{Z}_{12}^\dagger \mathbf{Z}_{22}^{-1} \mathbf{Z}_{12})^{-1}$ [18], which allows us to express the objective function for FD-DFE as

$$\text{OBJ} = \det(\mathbf{U}_{11}^{-1}) = \det(\mathbf{Z}_{11} - \mathbf{Z}_{12}^\dagger \mathbf{Z}_{22}^{-1} \mathbf{Z}_{12}). \quad (49)$$

2) *Upper Bound on Objective Function:* Unfortunately, the expression for OBJ in (49) depends on the FBF length N_{fb} , cf. (46), which is not desirable in practice. Additionally, due to the presence of \mathbf{Z}_{22}^{-1} in (49), it is also not straightforward to express the objective function in terms of \mathbf{A}_k and $\tilde{\mathbf{P}}_k$. To avoid these problems, we derive an upper bound for OBJ, which is independent of N_{fb} and directly related to the optimization variables.

Since matrix \mathbf{Z} in (17) is a positive semidefinite (PSD) matrix, \mathbf{Z}^{-1} , \mathbf{Z}_{22} , and \mathbf{U}_{11} are PSD matrices as well. Thus, \mathbf{U}_{11}^{-1} and $\mathbf{Z}_{12}^\dagger \mathbf{Z}_{22}^{-1} \mathbf{Z}_{12}$ are also PSD matrices. By exploiting the fact that $\det(\mathbf{A} + \mathbf{B}) \geq \det(\mathbf{A})$ if \mathbf{A} and \mathbf{B} are PSD matrices [18], we obtain

$$\det(\mathbf{U}_{11}^{-1}) \leq \det(\mathbf{Z}_{11}), \quad (50)$$

where equality holds i.f.f. $N_{fb} = 0$. Therefore, for the case of $N_{fb} = 0$, $\det(\mathbf{Z}_{11})$ is the exact value of OBJ. Otherwise, it is an upper bound for OBJ, which can be expressed as

$$\begin{aligned} \text{OBJ}_{ub} &= \det(\mathbf{Z}_{11}) = \det\left(\sum_{k=1}^{N_c} \Psi_k^{-1}\right) \\ &= \det\left(\sum_{k=1}^{N_c} \mathbf{V}_1 \hat{\Psi}_k^{-1} \mathbf{V}_1^\dagger\right) = \det\left(\sum_{k=1}^{N_c} \hat{\Psi}_k^{-1}\right), \end{aligned} \quad (51)$$

where we exploited $\det(\mathbf{V}_1) = 1$.

3) *Structures of Optimal Source and Relay Precoding Matrices:* Since we can always chose $\tilde{\mathbf{P}}_k$ such that $\sum_{k=1}^{N_c} \hat{\Psi}_k^{-1}$ is a diagonal matrix, cf. (45), the determinant in (51) is essentially the product of the diagonal entries of $\sum_{k=1}^{N_c} \hat{\Psi}_k^{-1}$. Consequently, OBJ_{ub} is equivalent to the objective function for the FD-LE receiver under the GMSE criterion. Thus, from Theorem 1 we obtain the following optimal structures for $\tilde{\mathbf{P}}_k$ and \mathbf{A}_k

$$\tilde{\mathbf{P}}_k^* = \bar{\mathbf{V}}_G^{(k)} \mathbf{\Lambda}_P^{(k)}, \quad \mathbf{A}_k^* = \bar{\mathbf{V}}_H^{(k)} \mathbf{\Lambda}_A^{(k)} \bar{\mathbf{U}}_G^{(k)\dagger}. \quad (52)$$

And the optimal \mathbf{P}_k^* is thus given by $\tilde{\mathbf{P}}_k^* \mathbf{V}_1^\dagger$. The remaining power allocation problem is identical to that for the GMSE

criterion for FD-LE, cf. (38). It is worth mentioning that for $N_{fb} > 0$, the upper bound OBJ_{ub} constitutes a tight approximation of the objective function OBJ as is illustrated in Section V.

D. Optimal Power Allocation

From the previous two subsections, it can be concluded that only two different types of power allocation problems have to be solved, namely the problems for the AMSE and GMSE criteria for FD-LE. The solutions to these problems are also applicable for the maxMSE criterion for FD-LE and all three criteria for FD-DFE. However, since the objective functions for the AMSE and GMSE criteria in (38) are not jointly convex w.r.t. the power allocation variables, the global optimal solution is difficult to obtain. Thus, in the following, we adopt a high signal-to-noise ratio (SNR) approximation for Φ_{km} , i.e., we assume $\sigma_u^2 \sigma_v^2$ is sufficient small such that it can be ignored in the denominator of (37), which leads to

$$\Phi_{km} \approx \tilde{\Phi}_{km} = \frac{P_{s,km} P_{r,km} g_{km}^2 h_{km}^2}{\sigma_v^2 P_{r,km} h_{km}^2 + \sigma_u^2 P_{s,km} g_{km}^2} + 1. \quad (53)$$

This approximation renders the optimization problem convex such that efficient methods can be applied for its solution. In particular, it can be shown that both objective functions $f_{\text{AMSE}}(\tilde{\Phi})$ and $f_{\text{GMSE}}(\tilde{\Phi})$, $\tilde{\Phi} = \{\tilde{\Phi}_{km}, \forall k, m\}$, are jointly convex w.r.t. the power allocation variables $P_{s,km}$ and $P_{r,km}$, respectively. The convexity of the objective functions can be proved straightforwardly by showing that the Hessian matrix w.r.t. $P_{s,km}$ and $P_{r,km}$ is positive semi-definite. We omit the proof here because of space constraints. Furthermore, all power constraints are affine in $P_{r,km}$ and $P_{s,km}$. Thus, the optimization problem in (38), with Φ_{km} approximated by $\tilde{\Phi}_{km}$, is a convex optimization problem.

We are now ready to derive an iterative power allocation algorithm. To this end, we introduce the Lagrangian of the considered power allocation problem

$$\begin{aligned} \mathcal{L} &= f_X(\tilde{\Phi}_{km}) + \lambda \left[\sum_{k,m} P_{s,km} - P_S \right] + \mu \left[\sum_{k,m} P_{r,km} - P_R \right] \\ &\quad - \sum_{k,m} [\beta_{km} P_{s,km} + \gamma_{km} P_{r,km}], \end{aligned} \quad (54)$$

where λ and μ are the Lagrange multipliers for the sum power constraints for source and relay, respectively, and β_{km} and γ_{km} are the Lagrange multipliers for the individual power constraints for source and relay, respectively. Applying the Karush-Kuhn-Tucker (KKT) conditions to (54), which are sufficient and necessary conditions for convex optimization problems [19], we obtain the optimal solution of the considered problem as

$$\begin{aligned} P_{s,km} &= \frac{\sigma_v^2 P_{r,km}^2 h_{km}^2}{g_{km}^2 (P_{r,km}^2 h_{km}^2 + \sigma_u^2)} \left(\sqrt{\frac{g_{km}^2}{\lambda (\ln 2) B_m \sigma_u^2 \sigma_v^2}} - 1 \right)^+ \\ P_{r,km} &= \frac{\sigma_u^2 P_{s,km}^2 g_{km}^2}{h_{km}^2 (P_{s,km}^2 g_{km}^2 + \sigma_v^2)} \left(\sqrt{\frac{h_{km}^2}{\mu (\ln 2) B_m \sigma_u^2 \sigma_v^2}} - 1 \right)^+, \end{aligned} \quad (55)$$

TABLE I

ALGORITHM FOR FINDING THE OPTIMAL POWER ALLOCATION. ϵ_1 AND ϵ_2 ARE SMALL CONSTANTS, E.G. $\epsilon_1 = \epsilon_2 = 10^{-4}$.

1	Initialize $\mu^{[1]}$ and $\lambda^{[1]}$
2	Initialize $P_{s,km}^{[1]}, P_{r,km}^{[1]}, \forall k, m$. Set $P_{s,km}^{rec} = P_{s,km}^{[1]}, P_{r,km}^{rec} = P_{r,km}^{[1]}, \forall k, m$.
3	Repeat Set iteration number to $n = 2$. Repeat for $m = 1 : M, k = 1 : N_c$ Find $P_{s,km}^{[n]}$ from (55) using $P_{r,km}^{rec}$ and $\lambda^{[n-1]}$. end for Update $\lambda^{[n]}$ using (56). $n = n + 1$. until $ \lambda^{[n+1]} - \lambda^{[n]} < \epsilon_1$, set $P_{s,km}^{rec} = P_{s,km}^{[n]}$. Set iteration number to $l = 2$. Repeat for $m = 1 : M, k = 1 : N_c$ Find $P_{r,km}^{[l]}$ from (55) using $P_{s,km}^{rec}$ and $\mu^{[l-1]}$. end for Update $\mu^{[l]}$ using (57). $l = l + 1$. until $ \mu^{[l+1]} - \mu^{[l]} < \epsilon_2$, set $P_{r,km}^{rec} = P_{r,km}^{[l]}$. until $P_{r,km}^{rec}$ and $P_{s,km}^{rec}$ converge.
4	$P_{s,km}^{rec}$ and $P_{r,km}^{rec}, \forall k, m$, are the optimal solution.

where $B_m = 1$ and $B_m = \sum_{k=1}^{N_c} (\tilde{\Phi}_{km} + 1)^{-1}$ for the AMSE and the GMSE criteria, respectively, and $[x]^+ = \max(0, x)$. The Lagrange multipliers λ and μ , which are chosen to satisfy the sum power constraint for source and relay, respectively, can be found with the following subgradient method [19]

$$\lambda^{[n+1]} = \left[\lambda^{[n]} - \epsilon_1 \left(\sum_{k=0}^{N_c-1} \sum_{m=1}^M P_{s,km} - P_S \right) \right]^+ \quad (56)$$

$$\mu^{[n+1]} = \left[\mu^{[n]} - \epsilon_2 \left(\sum_{k=0}^{N_c-1} \sum_{m=1}^M P_{r,km} - P_R \right) \right]^+, \quad (57)$$

where n is the iteration index, and $\epsilon_i, i = 1, 2$, are step sizes. From (55), we observe that the optimal $P_{s,km}$ depends on $P_{r,km}$ and vice versa. To tackle this problem, we propose the algorithm in Table I to iteratively find the optimal power allocations. Convergence of this algorithm to the optimal solution is guaranteed because of the convexity of the considered optimization problem. Note that if either $P_{s,km}$ or $P_{r,km}$ is equal to 0, the other variable will also be 0. This result is intuitively pleasing since, if for example the (m, k) th subchannel is shut down in the S - R link, there is no need to waste power on this subchannel in the R - D link. It is also worth noting that for the GMSE criterion, $P_{s,km}$ and $P_{r,km}$ are functions of $\tilde{\Phi}_{km}$, which means the optimal $P_{s,km}$ and $P_{r,km}$ for the k th frequency tone depend on the power allocations in all other frequency tones. Therefore, finding the optimal solution requires a higher complexity for the GMSE criterion than for the AMSE criterion.

E. Suboptimal Power Allocation Schemes

Since the proposed precoding matrix optimization scheme involves an iterative power allocation algorithm and considerable information exchange between source and relay, it is desirable to investigate suboptimal approaches with lower complexity and reduced feedback overhead. One option is to adopt equal power allocation at the source and to optimize only the power allocation at the relay. We refer to the corresponding

scheme as EPA-S. EPA-S eliminates the need for information exchange between source and relay for power allocation, hence guaranteeing faster convergence of the power allocation algorithm. However, the EPA-S scheme still requires CSI feedback from the relay to the source for computation of the source precoding matrix. In order to completely avoid CSI feedback, one can perform precoding at the relay only, which we refer to as ROP scheme. For FD-DFE, we also introduce the UPS scheme, which applies only the unitary precoding matrix \mathbf{V}_1 at the source. This is motivated by the result in Section IV-B, where it is shown that this unitary matrix can balance the MSEs of the different spatial streams. Similar to ROP, the UPS scheme has the advantage of a reduced feedback overhead compared to optimal power allocation and the EPA-S scheme as the source only needs to acquire knowledge of the $M \times M$ unitary matrix \mathbf{V}_0 .

V. SIMULATION RESULTS

In this section, we evaluate the performance of the proposed source and relay precoding schemes using simulations. We assume that each data block contains $N_c = 64$ symbols. The channels are modelled as uncorrelated Rayleigh block fading channels with power delay profile $p[n] = \frac{1}{\sigma_t} \sum_{l=0}^{L_x-1} e^{-n/\sigma_t} \delta[n-l]$ [20], where $L_x \in \{L_g, L_h\}$ and $\sigma_t = 2$, which corresponds to moderately frequency-selective fading. For convenience, we set the values of $L_g, L_h, N_{g,s}$, and $N_{g,r}$ all equal to 16. Unless stated otherwise, N_{fb} is set to 15. We assume identical noise variances for both links, i.e., $\sigma_u^2 = \sigma_v^2$, and define the source and relay SNRs as $(E_b/N_0)_s \triangleq \frac{P_S}{N_b N_s N_c \sigma_u^2}$ and $(E_b/N_0)_r \triangleq \frac{P_R}{N_b N_s N_c \sigma_v^2}$, respectively, where N_b is the number of bits per symbol. For all simulation results shown, we set $(E_b/N_0)_s = 16$ dB and examine bit error rate (BER) and capacity as functions of $(E_b/N_0)_r$. All simulations are averaged over at least 10,000 independent channel realizations. In the following, the proposed joint source and relay precoding design is referred to as JSR, and the notation $\{M, N_s, N_r, N_d\}$ is used to specify a system with the parameters appearing in the bracket.

A. Convergence of the Algorithm and Tightness of OBJ_{ub} for FD-DFE

In Figs. 2 and 3, we show the value of the objective function for the GMSE criterion versus the numbers of inner and outer iterations for different values of $(E_b/N_0)_r$ for a $\{2, 2, 2, 2\}$ MIMO relay system⁵. We define an outer iteration as one optimization of $\{P_{s,km}\}$ or $\{P_{r,km}\}$ in the algorithm shown in Table I, and the update of $\{P_{s,km}\}$ ($\{P_{r,km}\}$) in each outer iteration constitutes one inner iteration. The reference lines indicate the optimal values of the objective function. Figs. 2 and 3 show that it takes at most ten inner iterations to obtain an intermediate solution for $\{P_{s,km}\}$ or $\{P_{r,km}\}$, and at most three outer iterations to obtain the final solutions for $\{P_{s,km}\}$ and $\{P_{r,km}\}$. Also, from Fig. 3, we observe that the largest improvement of the objective function value is obtained in the first and second outer iterations when $(E_b/N_0)_r$ is small

⁵Similar results also hold for the AMSE criterion.

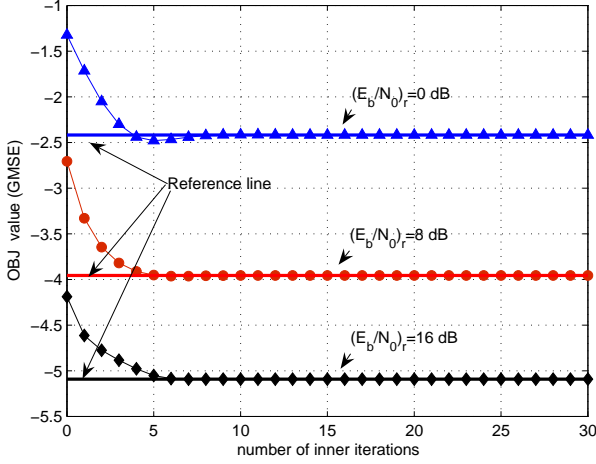


Fig. 2. Objective function value (GMSE criterion) versus number of inner iterations.

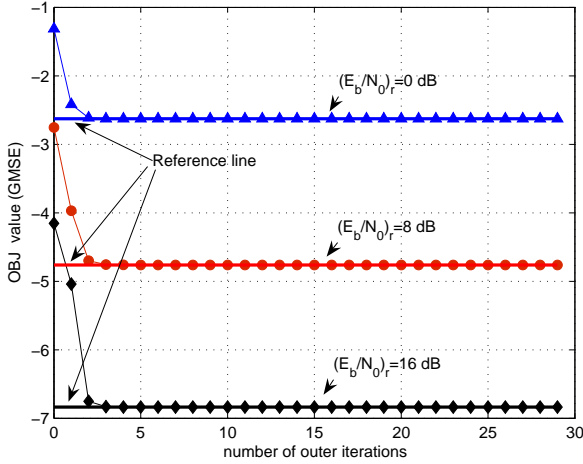


Fig. 3. Objective function value (GMSE criterion) versus number of outer iterations.

and large, respectively. This observation suggests that for low SNR, optimizing the source or the relay power is sufficient to realize most of the achievable performance gain, while for high SNR, a joint optimization of the source and relay powers is beneficial.

In Fig. 4, we show the values of the objective function, OBJ, for FD-DFE, cf. (49), for different values of N_{fb} . Note that OBJ for $N_{fb} = 0$ serves as the upper bound, OBJ_{ub} , for the general objective function. From the figure, we observe that the upper bound is very close to the objective function for all considered values of N_{fb} , especially for medium to high SNR and for the $\{2, 3, 3, 3\}$ system. Therefore, OBJ_{ub} constitutes a good approximation for the objective function for the FD-DFE receiver.

B. Comparison of SC-FDE and OFDM for JSR Precoding

In Fig. 5, we show the BER of uncoded quaternary phase-shift keying (QPSK) as a function of $(E_b/N_0)_r$ for the

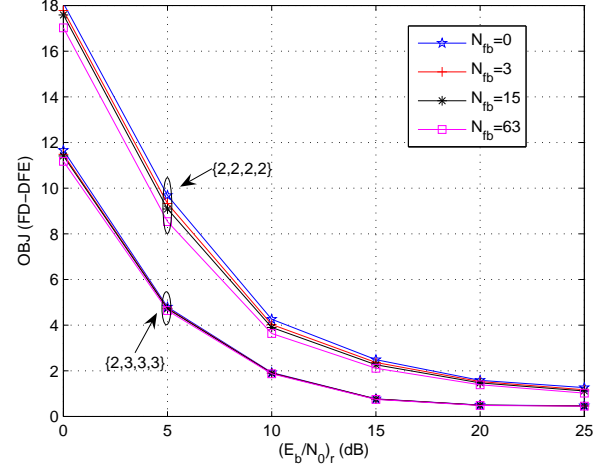


Fig. 4. Objective function value of the FD-DFE receiver for different values of N_{fb} .

proposed FD-LE based MIMO relay system for the three considered precoding matrix optimization criteria. For FD-DFE, only the GMSE criterion is considered as for FD-DFE all three criteria are equivalent. For comparison, the performance of a MIMO-OFDM relay system optimized under the same criteria is also included [6]. The figure shows that for the $\{2, 2, 2, 2\}$ system, the proposed MIMO relay system with an FD-LE receiver outperforms the corresponding OFDM-based system by a large margin since, in contrast to uncoded OFDM, FD-LE is able to exploit the frequency diversity offered by the channel. In addition, for both FD-LE and OFDM, the system employing the maxMSE criterion offers the best performance since the worst-case MSE is minimized. For FD-DFE, the performance improvement compared to FD-LE and OFDM is remarkable and a much higher diversity gain is observed. On the other hand, for the $\{2, 3, 3, 3\}$ system, we observe that the performance gaps between FD-DFE, FD-LE, and OFDM become smaller. Surprisingly, using the maxMSE criterion, the optimized OFDM and FD-LE systems achieve a performance very close to that of FD-DFE. This is due to fact that the additional antennas offer additional spatial diversity which helps OFDM and FD-LE to effectively avoid the deep spectrum nulls that otherwise negatively affect their performance in frequency-selective fading.

In Fig. 6, we investigate the capacity of the OFDM and SC-FDE systems with different optimization criteria. As expected, the systems optimized under the GMSE criterion have the best performance since minimizing the GMSE is equivalent to maximizing the capacity. In general, the capacity achieved by the considered MIMO-OFDM relay systems is higher than that of the corresponding FD-LE relay systems, except for the case when both systems are optimized based on the maxMSE criterion. Indeed, the OFDM system optimized under the maxMSE criterion suffers from the worst capacity performance among all the considered schemes since the available power is mainly used to improve the MSE of the subcarriers with bad channel conditions instead of taking advantage of the subcarriers with good channel conditions. In addition, Fig. 6 shows that for FD-

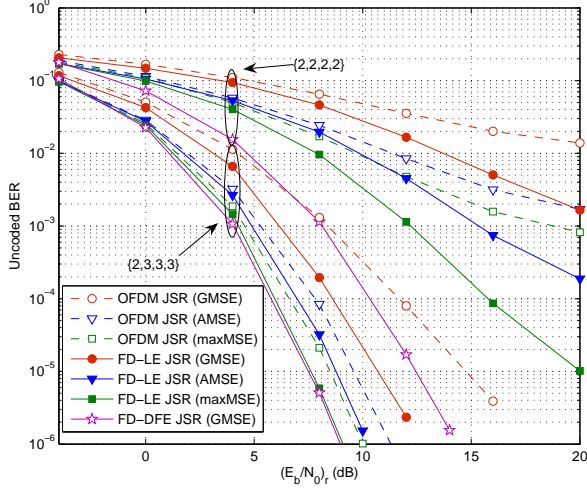


Fig. 5. Uncoded BER of $\{2, 2, 2, 2\}$ and $\{2, 3, 3, 3\}$ MIMO relay systems for JSR precoding using different optimization criteria.

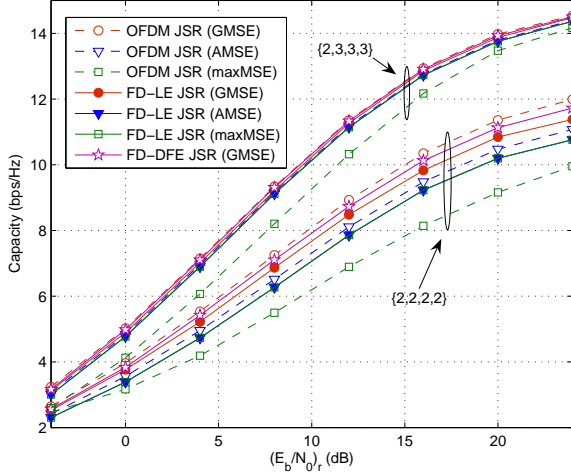


Fig. 6. Capacity of $\{2, 2, 2, 2\}$ and $\{2, 3, 3, 3\}$ MIMO relay systems with JSR relay precoding for different optimization criteria.

LE, the AMSE and maxMSE criteria lead to exactly the same capacity, which implies that the unitary rotation of the source precoding matrix does not influence the capacity of the system. Furthermore, the capacity achieved with FD-DFE is larger than that achieved with any of the FD-LE schemes and very close to that of OFDM. This is due to the lower stream MSEs of FD-DFE compared to FD-LE, which translates into larger stream SINRs and larger system capacity. For the $\{2, 3, 3, 3\}$ system, we observe that FD-LE and OFDM achieve almost the same performance for the AMSE and GMSE criteria, implying that with more source/relay/destination antennas, FD-LE will approach the achievable capacity of the OFDM system.

C. Performance of Suboptimal Power Allocation Schemes

In Figs. 7 and 8, we plot the uncoded and coded BERs for the suboptimal power allocation schemes discussed in Section IV-D for a $\{2, 2, 2, 2\}$ system, respectively. For the coded case,

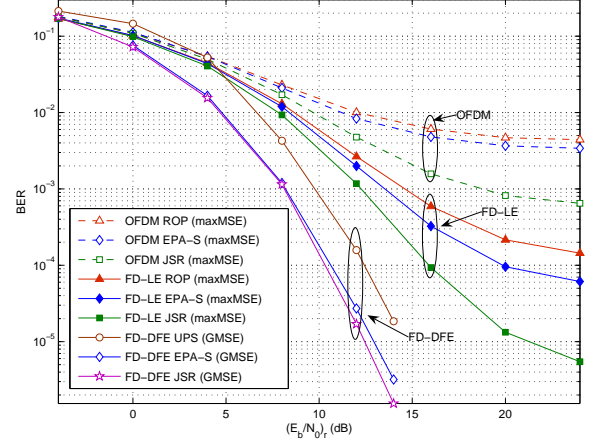


Fig. 7. Uncoded BER of a $\{2, 2, 2, 2\}$ MIMO relay system with JSR and suboptimal precoding schemes.

the standard rate-1/2 convolution code with generator matrix $(133, 171)_{oct}$ is adopted. The OFDM and FD-LE systems are both optimized under the maxMSE criterion. The FD-DFE system is optimized under the GMSE criterion since for FD-DFE all three considered criteria are equivalent to the GMSE criterion. From Fig. 7 we observe that for uncoded transmission, the FD-LE system outperforms the OFDM system if both employ the same precoding technique. Fig. 7 also shows that for FD-LE and OFDM, EPA-S and ROP suffer from a considerable performance degradation compared to JSR, while for FD-DFE, the performance loss is relatively small for UPS and almost negligible for EPA-S. By coding across subcarriers, OFDM-based relay systems can also exploit the frequency diversity of the channel and significantly improve their BER performance, cf. Fig. 8. Nevertheless, the coded FD-LE system still outperforms the OFDM system if the same precoding technique is assumed in both cases. Also, Fig. 8 reveals that channel coding significantly reduces the performance loss caused by suboptimal precoding techniques for both OFDM and FD-LE.

Since the performance of FD-DFE depends on the the number of FBF taps, in Fig. 9, we investigate the influence of N_{fb} on the performance of a $\{2, 2, 2, 2\}$ system. The results show that while the value of N_{fb} has limited impact on the performance of EPA-S, it does play a critical role for UPS. The reason is that for EPA-S, the equivalent S - R - D channel is diagonalized into M parallel channels, thus eliminating the inter-stream interference at the receiver. However, for the case of UPS, the equivalent end-to-end channel is not fully diagonalized and the received symbols experience inter-stream interference. Consequently, a FBF with sufficiently large N_{fb} is required to cancel out this interference. As can be inferred from Fig. 9, there is a complexity tradeoff between the transmitter and the receiver for FD-DFE. For EPA-S, since a small number of FBF taps (e.g., $N_{fb} = 3$) is sufficient to achieve good performance, the receiver complexity is similar to that of FD-LE. However, comparatively complex FD signal processing has to be carried out at the transmitter. This characteristic

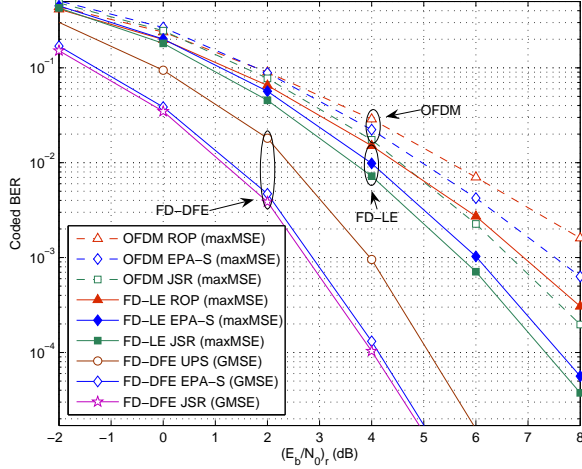


Fig. 8. Coded BER of a $\{2, 2, 2, 2\}$ MIMO relay system with JSR and suboptimal precoding schemes.

makes EPA-S suitable for the downlink transmission. For the UPS scheme, on the other hand, the transmit processing is very simple since the single tap precoding matrix \mathbf{V}_1 can be directly implemented in the TD. In addition, the feedback overhead is low as \mathbf{V}_1 is identical for all frequency tones. However, UPS requires a longer and thus more complex FBF to achieve a high performance. These characteristics make UPS a very promising scheme for uplink transmission.

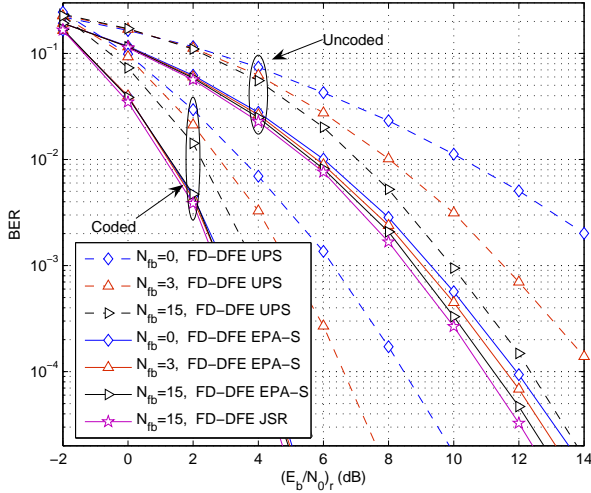


Fig. 9. Uncoded and coded BER of a $\{2, 2, 2, 2\}$ MIMO relay system with FD-DFE using JSR and suboptimal precoding schemes with different numbers of FBF taps.

VI. CONCLUSION

In this paper, we have tackled the problem of transceiver design for MIMO relay systems employing SC-FDE. The optimal MMSE FD-LE and FD-DFE filters at the destination were derived, and the optimal structures of the source and relay precoding matrices were obtained in closed form for a

general family of objective functions. For systems employing an FD-DFE receiver, we first showed that the considered objective functions are all equivalent and derived an upper bound on the original objective functions, which was shown to be equal to the GMSE objective function for the FD-LE receiver. The remaining power allocation problem was solved globally by using a high SNR approximation of the objective function and efficient convex optimization methods. Our results show that the proposed SC-FDE relaying schemes outperform the corresponding OFDM schemes in terms of both coded and uncoded BER. In addition, the performance gap between SC-FDE and OFDM relay systems decreases when the number of source/relay/destination antennas is larger than the number of data streams. Furthermore, we have shown that the proposed suboptimal power allocation schemes can reduce the system complexity and feedback overhead at the expense of a moderate performance degradation, especially in case of coded transmission, making them promising candidates for practical relay systems.

APPENDIX

We first provide some relevant definitions and lemmas that will be used in the proof.

Definition 1 [15,1.A.1]: Given two $N \times 1$ real vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$. Let $x_{[1]}, \dots, x_{[N]}$ and $y_{[1]}, \dots, y_{[N]}$ denote the components of \mathbf{x} and \mathbf{y} in decreasing order. Then, \mathbf{x} is majorized by \mathbf{y} , or $\mathbf{x} \prec \mathbf{y}$, if $\sum_{i=1}^k x_{[i]} \leq \sum_{i=1}^k y_{[i]}$ for $k < N$ and $\sum_{i=1}^N x_{[i]} = \sum_{i=1}^N y_{[i]}$. Vector \mathbf{x} is weakly majorized by \mathbf{y} , or $\mathbf{x} \prec_w \mathbf{y}$, if $\sum_{i=1}^k x_{[i]} \leq \sum_{i=1}^k y_{[i]}, \forall k$.

Definition 2 [15,3.A.1]: A real function f is Schur-convex if for $\mathbf{x} \prec \mathbf{y}$, we have $f(\mathbf{x}) \leq f(\mathbf{y})$. Similarly, f is Schur-concave if for $\mathbf{x} \prec \mathbf{y}$, we have $f(\mathbf{x}) \geq f(\mathbf{y})$.

Lemma 1 [15,9.B.1]: For a Hermitian matrix \mathbf{A} with $\text{diag}[\mathbf{A}]$ and $\lambda(\mathbf{A})$ denoting vectors containing the main diagonal elements and the eigenvalues of \mathbf{A} arranged in decreasing order, respectively, we have $\text{diag}[\mathbf{A}] \prec \lambda(\mathbf{A})$.

Lemma 2 [15,9.H.2]: For M matrices $\mathbf{A}_i \in \mathbb{C}^{N \times N}, i = 1, \dots, M$, let $\mathbf{B} = \mathbf{A}_1 \mathbf{A}_2 \dots \mathbf{A}_M$. Then, $\sigma(\mathbf{B}) \prec_w \sigma(\mathbf{A}_1) \odot \sigma(\mathbf{A}_2) \odot \dots \odot \sigma(\mathbf{A}_M)$, where $\sigma(\mathbf{X})$ denotes the vector containing the singular values of matrix \mathbf{X} arranged in decreasing order and \odot denotes the element-wise product of two vectors.

Lemma 3 [15,3.A.8]: A real function f satisfies $\mathbf{x} \prec_w \mathbf{y} \Rightarrow f(\mathbf{x}) \leq f(\mathbf{y})$ if and only if f is Schur-convex and increasing.

Lemma 4 [15,9.H.1]: For two Hermitian positive semidefinite matrices $\{\mathbf{A}, \mathbf{B}\} \in \mathbb{C}^{N \times N}$ with eigenvalues $\lambda_{\mathbf{A},i}, \lambda_{\mathbf{B},i}$, arranged in the same order, we have $\text{tr}(\mathbf{A}\mathbf{B}) \geq \sum_{i=1}^N \lambda_{\mathbf{A},i} \lambda_{\mathbf{B},N-i+1}$.

Lemma 5 [15, p.7]: For a vector $\mathbf{x} \in \mathbb{C}^{N \times 1}$, we have $\sum_{i=1}^N (x_i/N) \mathbf{1} \prec \mathbf{x}$, where $\mathbf{1}$ is the all-ones vector.

Lemma 6: For Hermitian matrices \mathbf{A}_k and $\hat{\mathbf{A}}_k, k = 1, 2, \dots, N_c$, if $\text{diag}[\mathbf{A}_k] \prec \text{diag}[\hat{\mathbf{A}}_k]$, we have $\text{diag}[\sum_{k=1}^{N_c} \mathbf{A}_k] \prec \text{diag}[\sum_{k=1}^{N_c} \hat{\mathbf{A}}_k]$.

Proof: According to Definition 1, if $\text{diag}[\mathbf{A}_k] \prec \text{diag}[\hat{\mathbf{A}}_k]$, we have $\sum_{i=1}^j a_{[ki]} \leq \sum_{i=1}^j \hat{a}_{[ki]}$ for $j < N$ and $\sum_{i=1}^N a_{[ki]} = \sum_{i=1}^N \hat{a}_{[ki]}$ for $j = N$, where $a_{[ki]}$ and $\hat{a}_{[ki]}$ are the i th diagonal entries of \mathbf{A}_k and $\hat{\mathbf{A}}_k$, respectively. Since $a_{[ki]}$ and $\hat{a}_{[ki]}, \forall k, i$, are non-negative, we have

$\sum_{k=1}^{N_c} \sum_{i=1}^j a_{[ki]} \leq \sum_{k=1}^{N_c} \sum_{i=1}^j \hat{a}_{[ki]}$ for $j < N$ and $\sum_{k=1}^{N_c} \sum_{i=1}^N a_{[ki]} = \sum_{k=1}^{N_c} \sum_{i=1}^N \hat{a}_{[ki]}$ for $j = N$. Therefore, we have $\sum_{k=1}^{N_c} \text{diag}[\mathbf{A}_k] \prec \sum_{k=1}^{N_c} \text{diag}[\hat{\mathbf{A}}_k]$, which is equivalent to $\text{diag}[\sum_{k=1}^{N_c} \mathbf{A}_k] \prec \text{diag}[\sum_{k=1}^{N_c} \hat{\mathbf{A}}_k]$.

Lemma 7 [15, 9.B.2]: For a diagonal matrix $\mathbf{D} \in \mathbb{C}^{M \times M}$, there is a unitary matrix \mathbf{U} such that $\mathbf{A} = \mathbf{U}^\dagger \mathbf{D} \mathbf{U}$ has identical diagonal entries equal to $\text{tr}(\mathbf{D})/M$.

We now set out to prove the optimal structure of the source and relay precoding matrices when $f(\text{diag}[\hat{\mathbf{E}}])$ is a Schur-concave increasing function w.r.t. $\text{diag}[\hat{\mathbf{E}}]$. Let us begin with the core term in the expression for $\hat{\mathbf{E}}$ in (12), which is given by

$$\Psi_k^{-1} = \mathbf{I}_M - \underbrace{\sigma_s^2 \mathbf{Q}_k^\dagger (\sigma_s^2 \mathbf{Q}_k \mathbf{Q}_k^\dagger + \sigma_v^2 \mathbf{H}_k \mathbf{A}_k \mathbf{A}_k^\dagger \mathbf{H}_k^\dagger + \sigma_u^2 \mathbf{I}_{N_d})^{-1} \mathbf{Q}_k}_{\mathbf{\Upsilon}_k}, \quad (58)$$

where we employed the matrix inversion lemma. The CV matrix can now be expressed as

$$\hat{\mathbf{E}} = \frac{\sigma_s^2}{N_c} \sum_{k=1}^{N_c} \Psi_k^{-1} = \frac{\sigma_s^2}{N_c} (N_c \mathbf{I}_M - \sum_{k=1}^{N_c} \mathbf{\Upsilon}_k). \quad (59)$$

Since $f(\text{diag}[\sigma_s^2 \mathbf{I}_M - \frac{\sigma_s^2}{N_c} \sum_{k=1}^{N_c} \mathbf{\Upsilon}_k])$ is a Schur-concave decreasing function w.r.t. $\text{diag}[\sum_{k=1}^{N_c} \mathbf{\Upsilon}_k]$, $-f(\text{diag}[\sigma_s^2 \mathbf{I}_M - \frac{\sigma_s^2}{N_c} \sum_{k=1}^{N_c} \mathbf{\Upsilon}_k])$ is a Schur-convex increasing function w.r.t. $\text{diag}[\sum_{k=1}^{N_c} \mathbf{\Upsilon}_k]$. $\mathbf{\Upsilon}_k$ can be further expressed as

$$\mathbf{\Upsilon}_k = \mathbf{X}_k^\dagger \mathbf{Y}_k^\dagger (\mathbf{J}_k \mathbf{J}_k^\dagger + \sigma_u^2 \mathbf{I}_{N_d})^{-1} \mathbf{Y}_k \mathbf{X}_k, \quad (60)$$

where we have used the definitions

$$\begin{aligned} \mathbf{X}_k &= \sigma_s \mathbf{G}_k \mathbf{P}_k, \quad \mathbf{Y}_k = \mathbf{H}_k \mathbf{A}_k, \\ \mathbf{J}_k &= \mathbf{Y}_k (\mathbf{X}_k \mathbf{X}_k^\dagger + \sigma_v^2 \mathbf{I}_{N_r})^{1/2}. \end{aligned} \quad (61)$$

By rewriting \mathbf{Y}_k in terms of \mathbf{X}_k and \mathbf{J}_k , we obtain $\mathbf{Y}_k = \mathbf{J}_k (\mathbf{X}_k \mathbf{X}_k^\dagger + \sigma_v^2 \mathbf{I}_{N_r})^{-1/2}$. Plugging this result into (60) leads to

$$\begin{aligned} \mathbf{\Upsilon}_k &= \mathbf{X}_k (\mathbf{X}_k \mathbf{X}_k^\dagger + \sigma_v^2 \mathbf{I}_{N_r})^{-1/2} \mathbf{J}_k^\dagger (\mathbf{J}_k \mathbf{J}_k^\dagger + \sigma_u^2 \mathbf{I}_{N_d})^{-1} \\ &\quad \mathbf{J}_k (\mathbf{X}_k \mathbf{X}_k^\dagger + \sigma_v^2 \mathbf{I}_{N_r})^{-1/2} \mathbf{X}_k. \end{aligned} \quad (62)$$

Now, we define the following SVDs

$$\begin{aligned} \mathbf{X}_k &= \mathbf{U}_X^{(k)} \mathbf{\Lambda}_X^{(k)} \mathbf{V}_X^{(k)}, \quad \mathbf{J}_k = \mathbf{U}_J^{(k)} \mathbf{\Lambda}_J^{(k)} \mathbf{V}_J^{(k)}, \\ (\mathbf{X}_k \mathbf{X}_k^\dagger + \sigma_v^2 \mathbf{I}_{N_r})^{-1/2} &= \mathbf{U}_X^{(k)} (\mathbf{\Lambda}_X^{(k)2} + \sigma_v^2 \mathbf{I}_M)^{-1/2} \mathbf{\Omega}^{(k)}, \end{aligned} \quad (63)$$

where $\mathbf{U}_X^{(k)} \in \mathbb{C}^{N_r \times M}$, $\mathbf{U}_J^{(k)} \in \mathbb{C}^{N_d \times M}$, $\{\mathbf{V}_J^{(k)}, \mathbf{\Omega}^{(k)}\} \in \mathbb{C}^{M \times N_r}$, $\{\mathbf{\Lambda}_X^{(k)}, \mathbf{V}_X^{(k)}, \mathbf{\Lambda}_J^{(k)}\} \in \mathbb{C}^{M \times M}$. The diagonal entries of $\mathbf{\Lambda}_X^{(k)}$ and $\mathbf{\Lambda}_J^{(k)}$ are both sorted in increasing order. We can use (63) to rewrite (62) as

$$\begin{aligned} \mathbf{\Upsilon}_k &= \mathbf{V}_X^{(k)\dagger} \mathbf{\Lambda}_X^{(k)} \mathbf{Q}_2^{(k)\dagger} (\mathbf{\Lambda}_X^{(k)2} + \sigma_v^2 \mathbf{I}_M)^{-1/2} \mathbf{Q}_1^{(k)\dagger} \\ &\quad \times (\mathbf{I}_M + \sigma_u^2 \mathbf{\Lambda}_J^{(k)-2})^{-1} \mathbf{Q}_1^{(k)} (\mathbf{\Lambda}_X^{(k)2} + \sigma_v^2 \mathbf{I}_M)^{-1/2} \\ &\quad \times \mathbf{Q}_2^{(k)} \mathbf{\Lambda}_X^{(k)} \mathbf{V}_X^{(k)}, \end{aligned} \quad (64)$$

where $\mathbf{Q}_1^{(k)} = \mathbf{V}_J^{(k)} \mathbf{U}_X^{(k)}$ and $\mathbf{Q}_2^{(k)} = \mathbf{\Omega}^{(k)} \mathbf{U}_X^{(k)}$. By applying Lemmas 1 and 2, we obtain

$$\begin{aligned} \text{diag}[\mathbf{\Upsilon}_k] &\prec \lambda(\mathbf{\Upsilon}_k) \\ &\prec_w \text{diag}[\underbrace{(\mathbf{I}_M + \sigma_v^2 \mathbf{\Lambda}_X^{(k)-2})^{-1}}_{\mathbf{D}_1^{(k)}} \underbrace{(\mathbf{I}_M + \sigma_u^2 \mathbf{\Lambda}_J^{(k)-2})^{-1}}_{\mathbf{D}_2^{(k)}}]. \end{aligned} \quad (65)$$

Therefore, $\text{diag}[\mathbf{\Upsilon}_k]$ is majorized by $\text{diag}[\mathbf{D}_1^{(k)} \mathbf{D}_2^{(k)}]$ when $\mathbf{V}_X^{(k)} = \mathbf{\Xi}_1$, $\mathbf{Q}_2^{(k)} = \mathbf{\Xi}_2$, $\mathbf{Q}_1^{(k)} = \mathbf{\Xi}_3$, where $\mathbf{\Xi}_i \in \mathbb{C}^{M \times M}$, $\forall i$, are arbitrary diagonal matrices with unit norm diagonal elements. Without loss of generality, we can choose $\mathbf{\Xi}_i = \mathbf{I}_M$, $\forall i$. Hence, we have

$$\mathbf{V}_X^{(k)} = \mathbf{I}_M, \quad \mathbf{\Omega}^{(k)} = \mathbf{U}_X^{(k)\dagger}, \quad \mathbf{V}_J^{(k)} = \mathbf{U}_X^{(k)\dagger}. \quad (66)$$

According to Lemma 6,

$$\text{diag}[\sum_{k=1}^{N_c} \mathbf{\Upsilon}_k] \prec_w \text{diag}[\sum_{k=1}^{N_c} (\mathbf{D}_1^{(k)} \mathbf{D}_2^{(k)})] \quad (67)$$

holds. From the fact that $-f(\text{diag}[\sigma_s^2 \mathbf{I}_M - \frac{\sigma_s^2}{N_c} \sum_{k=1}^{N_c} \mathbf{\Upsilon}_k])$ is a Schur-convex increasing function w.r.t. $\text{diag}[\sum_{k=1}^{N_c} \mathbf{\Upsilon}_k]$, (67), and Definition 2, we deduce that $-f(\text{diag}[\sigma_s^2 \mathbf{I}_M - \frac{\sigma_s^2}{N_c} \sum_{k=1}^{N_c} \mathbf{\Upsilon}_k]) \leq -f(\text{diag}[\sigma_s^2 \mathbf{I}_M - \frac{\sigma_s^2}{N_c} \sum_{k=1}^{N_c} \mathbf{D}_1^{(k)} \mathbf{D}_2^{(k)}])$, which is equivalent to

$$\begin{aligned} &f(\text{diag}[\sigma_s^2 \mathbf{I}_M - \frac{\sigma_s^2}{N_c} \sum_{k=1}^{N_c} \mathbf{\Upsilon}_k]) \\ &\geq f(\text{diag}[\sigma_s^2 \mathbf{I}_M - \frac{\sigma_s^2}{N_c} \sum_{k=1}^{N_c} \mathbf{D}_1^{(k)} \mathbf{D}_2^{(k)}]). \end{aligned} \quad (68)$$

Therefore, with the help of the matrices in (66), the value of the objective function can be reduced to $f(\text{diag}[\sigma_s^2 \mathbf{I}_M - \frac{\sigma_s^2}{N_c} \sum_{k=1}^{N_c} \mathbf{D}_1^{(k)} \mathbf{D}_2^{(k)}])$.

In the following, we derive the structure of the optimal source and relay precoding matrices minimizing the power consumption of the source and the relay. For simplicity of notation, we only consider the case when $N_s = N_r = N_d$. The proof can be easily extended to the case where N_s , N_r , and N_d have different values. From (63), we have $\sigma_s \mathbf{U}_G^{(k)} \mathbf{\Lambda}_G^{(k)} \mathbf{V}_G^{(k)\dagger} \mathbf{P}_k = \mathbf{U}_X^{(k)} \mathbf{\Lambda}_X^{(k)} \mathbf{V}_X^{(k)}$, which can be used to express the source power consumption at frequency tone k as

$$\begin{aligned} &\sigma_s^2 \text{tr}(\mathbf{P}_k \mathbf{P}_k^\dagger) \\ &= \text{tr}(\underbrace{\mathbf{\Lambda}_G^{(k)-1} \mathbf{U}_G^{(k)\dagger} \mathbf{U}_X^{(k)}}_{\mathbf{Q}_3^{(k)\dagger}} \underbrace{\mathbf{\Lambda}_X^{(k)2} \mathbf{U}_X^{(k)\dagger} \mathbf{U}_G^{(k)}}_{\mathbf{Q}_3^{(k)}} \mathbf{\Lambda}_G^{(k)-1}) \\ &\geq \text{tr}(\bar{\mathbf{\Lambda}}_G^{(k)-2} \mathbf{\Lambda}_X^{(k)2}), \end{aligned} \quad (69)$$

where the inequality follows from Lemma 4, and the diagonal matrix $\bar{\mathbf{\Lambda}}_G^{(k)} \in \mathbb{C}^{M \times M}$ contains the M largest singular values of \mathbf{G}_k . Therefore, in order to minimize the source power, we need to choose $\mathbf{Q}_3^{(k)} = [\mathbf{I}_M \quad \mathbf{0}_{M \times (N_s - M)}]^T$, i.e., $\mathbf{U}_X^{(k)} = \bar{\mathbf{U}}_G^{(k)}$, where $\bar{\mathbf{U}}_G^{(k)}$ contains the M right-most columns of $\mathbf{U}_G^{(k)}$.

Recalling from (69) that $\mathbf{V}_X^{(k)} = \mathbf{I}_M$, the source matrix can be expressed as

$$\mathbf{P}_k = \frac{1}{\sigma_s} \bar{\mathbf{V}}_G^{(k)} \bar{\Lambda}_G^{(k)-1} \Lambda_X^{(k)} = \bar{\mathbf{V}}_G^{(k)} \Lambda_P^{(k)}, \quad (70)$$

where $\bar{\mathbf{V}}_G^{(k)}$ contains the M right-most columns of $\mathbf{V}_G^{(k)}$ and $\Lambda_P^{(k)} \triangleq \frac{1}{\sigma_s} \bar{\Lambda}_G^{(k)-1} \Lambda_X^{(k)}$. Next, from (63) we obtain $\mathbf{U}_H^{(k)} \Lambda_H^{(k)} \mathbf{V}_H^{(k)\dagger} \mathbf{A}_k = \mathbf{J}_k (\mathbf{X}_k \mathbf{X}_k^\dagger + \sigma_v^2 \mathbf{I}_{N_d})^{-1/2}$. Using this result, we can express the relay power consumption at frequency tone k as

$$\begin{aligned} & \text{tr}(\mathbf{A}_k (\mathbf{X}_k \mathbf{X}_k^\dagger + \sigma_v^2 \mathbf{I}_{N_r}) \mathbf{A}_k^\dagger) \\ &= \text{tr}(\Lambda_H^{(k)-1} \mathbf{U}_H^{(k)\dagger} \mathbf{J}_k \mathbf{J}_k^\dagger \mathbf{U}_H^{(k)} \Lambda_H^{(k)-1}) \\ &= \text{tr}(\Lambda_H^{(k)-1} \underbrace{\mathbf{U}_H^{(k)\dagger} \mathbf{U}_J^{(k)}}_{\mathbf{Q}_4^{(k)\dagger}} \Lambda_J^{(k)2} \underbrace{\mathbf{U}_J^{(k)\dagger} \mathbf{U}_H^{(k)}}_{\mathbf{Q}_4^{(k)}} \Lambda_H^{(k)-1}) \\ &\geq \text{tr}(\bar{\Lambda}_H^{(k)-2} \Lambda_J^{(k)2}), \end{aligned} \quad (71)$$

where the inequality follows from Lemma 4, and the diagonal matrix $\bar{\Lambda}_H^{(k)} \in \mathbb{C}^{M \times M}$ contains the M largest singular values of \mathbf{H}_k . In (71), equality holds for $\mathbf{Q}_4^{(k)} = [\mathbf{I}_M \mathbf{0}_{M \times (N_r - M)}]^T$, i.e., $\mathbf{U}_J^{(k)} = \bar{\mathbf{U}}_H^{(k)}$, where $\bar{\mathbf{U}}_H^{(k)}$ contains the M right-most columns of $\mathbf{U}_H^{(k)}$. Recalling from (66) that $\mathbf{U}_X^{(k)\dagger} = \Omega^{(k)} = \mathbf{V}_J^{(k)}$, and from (69) that $\mathbf{U}_X^{(k)} = \bar{\mathbf{U}}_G^{(k)}$, we obtain for \mathbf{A}_k the expression

$$\begin{aligned} \mathbf{A}_k &= \bar{\mathbf{V}}_H^{(k)} \bar{\Lambda}_H^{(k)-1} \Lambda_J^{(k)} (\Lambda_X^{(k)2} + \sigma_v^2 \mathbf{I}_M)^{-1/2} \bar{\mathbf{U}}_G^{(k)\dagger} \\ &= \bar{\mathbf{V}}_H^{(k)} \Lambda_A^{(k)} \bar{\mathbf{U}}_G^{(k)\dagger}, \end{aligned} \quad (72)$$

where $\bar{\mathbf{V}}_H^{(k)}$ and $\bar{\mathbf{U}}_G^{(k)}$ contain the M right-most columns of $\mathbf{V}_H^{(k)}$ and $\mathbf{U}_G^{(k)}$, respectively, and $\Lambda_A^{(k)} \triangleq \bar{\Lambda}_H^{(k)-1} \Lambda_J^{(k)} (\Lambda_X^{(k)2} + \sigma_v^2 \mathbf{I}_M)^{-1/2}$. Hence, we have proved that the expressions for the source and relay precoding matrices given in (70) and (72) minimize the objective function $f(\text{diag}[\hat{\mathbf{E}}])$, cf. (68), as well as the power consumption at the source and the relay, cf. (69) and (71).

Now, we turn our attention to the case when $f(\text{diag}[\hat{\mathbf{E}}])$ is a Schur-convex increasing function w.r.t. $\text{diag}[\hat{\mathbf{E}}]$. From Lemma 5 we know that $\frac{1}{M} \text{tr}(\hat{\mathbf{E}}) \mathbf{1} \prec \text{diag}[\hat{\mathbf{E}}]$. Combining this fact and Definition 2, we obtain the inequality

$$f(\text{diag}[\hat{\mathbf{E}}]) \geq f\left(\frac{1}{M} \text{tr}(\hat{\mathbf{E}}) \mathbf{1}\right), \quad (73)$$

where equality holds when the diagonal entries of $\hat{\mathbf{E}}$ are all equal to $\frac{1}{M} \text{tr}(\hat{\mathbf{E}})$. In the following, we show that by applying a unitary rotation to the source precoding matrix, we can achieve this equality. Using the eigenvalue decomposition $\mathbf{Q}_k^\dagger \mathbf{K}_k^{-1} \mathbf{Q}_k = \mathbf{U}_E^{(k)} \Lambda_E^{(k)} \mathbf{U}_E^{(k)\dagger}$, we can write $\Psi_k^{-1} = \mathbf{U}_E^{(k)} (\mathbf{I}_M + \Lambda_E^{(k)})^{-1} \mathbf{U}_E^{(k)\dagger}$. Let us consider the feasible source precoding matrix $\bar{\mathbf{P}}_k = \mathbf{P}_k \mathbf{U}_E^{(k)} \mathbf{V}_0$, where \mathbf{V}_0 is a unitary matrix and thus does not affect the power constraints. Replacing \mathbf{P}_k with $\bar{\mathbf{P}}_k$ in Ψ_k^{-1} , we arrive at

$$\Psi_k^{-1} = \mathbf{V}_0^\dagger (\mathbf{I}_M + \Lambda_E^{(k)})^{-1} \mathbf{V}_0, \quad (74)$$

which allows us to express the error CV matrix as

$$\hat{\mathbf{E}} = \frac{\sigma_s^2}{N_c} \sum_{k=1}^{N_c} \Psi_k^{-1} = \frac{\sigma_s^2}{N_c} \mathbf{V}_0^\dagger \sum_{k=1}^{N_c} (\mathbf{I}_M + \Lambda_E^{(k)})^{-1} \mathbf{V}_0. \quad (75)$$

Since $\sum_{k=1}^{N_c} (\mathbf{I}_M + \Lambda_E^{(k)})^{-1}$ is the sum of N_c diagonal matrices, it is also a diagonal matrix. Based on Lemma 7, we conclude that there exists a unitary matrix \mathbf{V}_0 such that $\hat{\mathbf{E}}$ has identical diagonal elements given by $\frac{1}{M} \text{tr}(\hat{\mathbf{E}})$. Since the objective function is an increasing function w.r.t. its arguments, minimizing the original Schur-convex objective function is now equivalent to minimizing $\frac{1}{M} \text{tr}(\hat{\mathbf{E}})$, which is a Schur-concave function. Therefore, the optimal structures of \mathbf{P}_k and \mathbf{A}_k are given by (70) and (72), respectively. Furthermore, as the resulting $\mathbf{U}_E^{(k)}$ can be shown to be an identity matrix, the source precoding matrix for Schur-convex functions is given by $\bar{\mathbf{P}}_k = \mathbf{P}_k \mathbf{V}_0 = \bar{\mathbf{V}}_G^{(k)} \Lambda_P^{(k)} \mathbf{V}_0$.

REFERENCES

- [1] R. Nabar, H. Bölcskei, and F. Kneubühler, "Fading relay channels: Performance limits and space-time signal design," *IEEE J. Sel. Areas Commun.*, vol. 22, pp. 1099–1109, 2004.
- [2] I. Hammerstrom and A. Wittneben, "Power allocation schemes for amplify-and-forward MIMO-OFDM relay links," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 2798–2802, 2007.
- [3] R. Mo and Y. Chew, "Precoder design for non-regenerative MIMO relay systems," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 5041–5049, 2007.
- [4] F. S. Tseng, W. R. Wu and J. Y. Wu, "Joint source/relay precoder design in nonregenerative cooperative systems using an MMSE criterion," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 4928–4933, 2009.
- [5] R. Mo and Y. Chew, "MMSE-based joint source and relay precoding design for amplify-and-forward MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 4668–4676, 2009.
- [6] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear nonregenerative multicarrier MIMO relay communication systems," *IEEE Trans. Signal Process.*, vol. 57, pp. 4837–4851, 2009.
- [7] Y. Rong and Y. Hua, "Optimality of diagonalization of multi-hop MIMO relays," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 6068–6077, 2009.
- [8] Y. Rong, "Optimal linear non-regenerative multi-hop MIMO relays with MMSE-DFE receiver at the destination," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 2268–2279, 2010.
- [9] C. Xing, S. Ma, Y. C. Wu, and T. S. Ng, "Transceiver design for dual-hop nonregenerative MIMO-OFDM relay systems under channel uncertainties," *IEEE Trans. Signal Process.*, vol. 58, pp. 6325–6339, 2010.
- [10] C. Jeong, B. Seo, S. Lee, H. Kim, and I. Kim, "Relay precoding for non-regenerative MIMO relay systems with partial CSI feedback," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 1698–1711, 2012.
- [11] D. Falconer, S. Ariyavisitkul, A. Benyamini-Seeyar, and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Commun. Mag.*, vol. 40, pp. 58–66, 2002.
- [12] F. Pancaldi, G. Vitetta, R. Kalbasi, N. Al-Dhahir, M. Uysal, and H. Mheidat, "Single-carrier frequency domain equalization," *IEEE Signal Process. Mag.*, vol. 25, pp. 37–56, 2008.
- [13] N. Al-Dhahir, and A. H. Sayed, "The finite-length multi-input multi-output MMSE-DFE," *IEEE Trans. Signal Process.*, vol. 48, pp. 2921–2936, 2000.
- [14] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, pp. 2381–2401, 2003.
- [15] A. W. Marshall, I. Olkin, and B. Arnold, *Inequalities: theory of majorization and its applications*. Springer Verlag, 2010.
- [16] J. K. Zhang, A. Kavcic, and K. M. Wong, "Equal-diagonal QR decomposition and its application to precoder design for successive-cancellation detection," *IEEE Trans. Information Theory*, vol. 51, pp. 154–172, 2005.
- [17] Y. Jiang, J. Li, and W. W. Hager, "Joint transceiver design for MIMO communications using geometric mean decomposition," *IEEE Trans. Signal Process.*, vol. 53, pp. 3791–3803, 2005.
- [18] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.
- [19] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [20] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Prentice Hall, 2002.

